Unethical Customer Behavior – Causes, Consequences, Detection and Managerial Implications

Dissertation
submitted to the
Faculty of Business, Economics and Informatics
of the University of Zurich

to obtain the degree of
Doktor der Wirtschaftswissenschaften, Dr. oec.
(corresponds to Doctor of Philosophy, PhD)

presented by

Zhao Yang
from China

approved in February 2017 at the request of
Prof. Dr. René Algesheimer
Prof. Dr. Martin Natter
Prof. Dr. Claudio J. Tessone

The Faculty of Business, Economics and Informatics of the University of Zurich hereby authorizes the printing of this dissertation, without indicating an opinion of the views expressed in the work.

Zurich, 15.02.2017

The Chairman of the Doctoral Board: Prof. Dr. Steven Ongena

To my parents Qigui and Qin, my wife Xue, and my son Chuanshuo:

This dissertation would never have been possible without your love.

Table of Contents

List of Tables

List of Figures

1 Introduction

Fraudulent behavior, also known as cheating, dishonesty and unethical behavior, has become a major phenomenon representing substantial amounts of losses. It is an action that falls outside of what is considered morally right or proper and it can happen among individuals, businesses, professionals, politicians, and governments [19]. Volkswagen, for instance, was in the news for unethical behavior on a massive scale. According to Ethics Alarms [7]: "Volkswagen, which manufactures many of the beasts, devised and installed a code functioning as a "defeat device" to sense when one of its diesel vehicles was being tested for nitrogen oxide emissions test. Once a test was detected, the software would reduce torque and NOx emissions, while under normal conditions, that is, when the vehicle was not being tested for emissions, the car would be guided by a separate program that would increase acceleration, torque, and fuel economy." Another well-known example could be "Banking scandals". UBS, the Union Bank of Switzerland, had a series of missteps in the recent past. Its Chief Executive, Oswald Grubel, resigned in November 2011 as the bank faced a loss of $2.3 billion due to a low-level rogue trader conducting unauthorized trade by sidestepping the bank's internal control mechanism and as the bank had to pay $45 million in fines for having inadequate internal control systems, which enabled the rogue trader to cheat [2].

Apart from these "famous" examples of unethical behavior, unethical customer behavior, which is defined as disobeying the moral principles and standards that guide behavior of individuals or groups as they obtain, use, and dispose of goods and services [11], is also an important aspect of unethical behavior in business. Common instances of unethical customer behavior includes theft, price-tag swapping, use of items without purchase, and de-shopping, etc. [18] Just based on these names, it's not surprising that scholars and practitioners focus mainly on negative aspects of unethical customer behavior. For instance, an earlier study [1] claimed that "You will regret letting the wrongdoer off the hook!" and it sought to help managers and executives by answering the following issues: (1) Where does bad behavior come from? (2) How to manage the unethical consumer, and (3) What moderates unethical actions?

This dissertation, however, points out that as bad character is not always tied to bad behavior, the consequences of unethical customer behavior are not certainly to be negative. Therefore, it's necessary to shift the focus away from the rightness or wrongness of the customers' behavior and to consider how retailers should respond to customers' unethical behavior in more comprehensive ways in order to benefit both customers and retailers.

The three independent papers in this dissertation are as follows.

The first paper is an empirical study. We employ longitudinal data from a Swiss online retailer that was founded in late 2011. The data contains various perspectives of its customers, e.g. their login activities, trading activities, purchasing activities, etc. Besides, we also receive a list of customers who violate the policy of the retailer, i.e. behaving unethically. We use the vector autoregressive models to reveal the connections between normal customers and fraudulent ones, and further study the impact of fraudulent behavior on the retailer. Based on the empirical analyses, we deliver generalizable results and provide managerial implications to the existing marketing research and practices.

The second paper is a review study. We provide an overview of existing literatures in standard economics, psychology, behavior economics, and neuroscience on unethical behavior in order to gain a deeper understanding of its causes and consequences. We then especially highlight the studies related to ordinary unethical behavior, no matter whether it is intentional or unintentional. After that, we briefly review different statistical fraud detections and elaborate the application of network science, e.g. social network analysis, and community detection algorithms, in fraud detection. In the end, we provide managerial recommendations to guide managers in real business.

The third paper is a methodological study. In this work, we have employed the Lancichinetti-Fortunato-Radicchi benchmark graphs to test eight state-of-the-art community detection algorithms [10, 9, 8, 5, 16, 14, 12, 3, 15, 13]. We quantify the accuracy of various community detection methods by using complementary measures. Based on simple network properties, we provide guidelines that help to choose the most adequate community detection algorithm for a given network. Community detection methods have drawn a lot of attentions

in fraud detection due to the fact that they can be helpful in discovering groups of structurally connected individuals. These methods, as the core of "community-based" anomaly detection techniques, can detect the graph objects that are rare and differ significantly from the majority of the reference objects in static and evolutionary networks in an unsupervised way [4, 17]. These graph objects can be nodes, edges, substructures, and the patterns of interactions.

1.1 Summary of Study One (Chapter 2)

In the first study, we probe the boundaries of the widely-held view about unethical customer behavior in two ways. First, we develop a more nuanced and broader picture of unethical customer behavior by postulating that not all such behaviors are harmful. Some unethical actions can have beneficial consequences. Second, we develop a theoretical framework of retailer response that accommodates this nuanced view of unethical customer behavior and provides retailers with a wider assortment of response options.

We analyze longitudinal data from a Swiss online retailer over 17 months using vector autoregressive models. We conclude that our research, together with our empirical analysis, compellingly uncovers the counter-intuitive phenomenon of unethical consumer behavior having predominantly positive consequences for the retailer and for other customers. It empirically illustrates the observation made by Donaldson and Dunfee (1994, p.258) [6], that "the ethical norms must be contoured to the rules of the specific economic practices and the notions of fairness of participants." Consequently, there is a strong need to consider unethical consumer behaviors in retailing contexts in more nuanced and balanced ways. This allows to devise solutions that are equally nuanced and lead to the best possible outcomes for customers and for retailers as a whole, even when such actions do not fall strictly within the parameters set forth by the influential moral philosophies.

This study is currently under the second round review at Journal of Retailing.

A brief overview of the this study is presented in Table 1, in which the research questions, contributions, data, methods, and results are summarized.

1.2 Summary of Study Two (Chapter 3)

In the second study, we first review various articles in standard economics, psychology, behavior economics, and neuroscience to understand the causes and consequences of unethical behavior. We then provide an overview of recent development in studying ordinary unethical behavior. After that we briefly review the statistical fraud detection techniques and highlight the application of network science in fraud detection.

Although from the standard economics perceptive committing fraud does not have any internal cost, individuals do have internal costs while committing fraud as the evidences from psychology and behavior economics suggest. These costs emerge because the actions an individual undertakes, while engaging in cheating behavior, are inconsistent with the internal ethical standards and moral principles. A growing body of research points to the fact that not only there exist internal psychological costs while cheating, but also that the costs are different among individuals. Therefore, as every individual might have very different internal costs, there exists no universal formula to predict precisely whether an individual commits fraud or not in a certain situation. Moreover, due to self-serving biases and bounded ethicality, people are sometimes not even aware of crossing ethical borders, and hence there are not many possibilities to reduce such unethical behavior. We believe that further development in neuroscience could shed light on understanding the underlying mechanisms behind these internal psychological costs, and will help us to prevent, detect, and punish fraudulent behavior more efficiently. A brief overview of the this study is presented in Table 2, in which the research questions, contributions, data, and results are summarized.

1.3 Summary of Study Three (Chapter 4)

In the third study, we test eight state-of-the-art community detection algorithms on the Lancichinetti-Fortunato-Radicchi benchmark graphs. We quantify the effects of the mixing parameter and the network size on accuracy and computing time of community detection algorithms. Our benchmark graphs have network sizes between 233 and 31,948 nodes, mixing

parameters between 0.03 and 0.75, and a fixed average degree of 20.

We conclude that by taking both accuracy and computing time into account, the "Multilevel" algorithm outperforms all the other algorithms on the set of benchmarks we have examined. We further provide guidelines that help to choose the most adequate community detection algorithm for a given network: For small networks, the community detection algorithms should be chosen based on their accuracies. Among all the algorithms, Infomap, Label propagation, Multilevel, Walktrap, Spinglass, and Edge betweenness algorithms are able to successfully uncover the structure of small networks when the mixing parameter is small. With increasing value of the mixing parameter, Infomap, Label propagation, and Edge betweenness algorithms are no longer suitable. For large networks, the algorithms should first be able to detect the organization of nodes in a reasonable time, and then have good accuracies. In this case, Infomap, Label propagation, Multilevel, and Walktrap algorithms are the a priori choices. After that, by taking the accuracy into account, Multilevel is superior to the other algorithms as it displays a performance drop for a larger value of the mixing parameter. We also point out that Spinglass and Multilevel algorithms can be used to get a rough idea about the value of the mixing parameter, which is usually unobservable. Limited by the computing time required, Spinglass algorithm cannot be applied on large networks.

This study has been published at Scientific Reports: Yang, Z., Algesheimer, R., & Tessone, C. J. (2016). A Comparative Analysis of Community Detection Algorithms on Artificial Networks. *Scientific Reports*, 6; doi: 10.1038/srep30750.

A brief overview of the this study is presented in Table 3, in which the research questions, contributions, data, methods, and results are summarized.

Table 1: Overview of the first study in the thesis.

Study 1: *When Unethical Customers Have Beneficial Effects: A Theory of Retailer Response to Unethical Customer Behaviors*		
Research questions What are the consequences of customer unethical behavior to the other customers, as well as to the retailer? How should the retailer response to unethical customer behaviors?	**Core contributions** Investigating the effect of unethical customer behavior to the other customers and the retailer. Developing a theoretical framework of retailer response that accommodates a more nuanced view of unethical customer behavior.	**Data basis** Daily customer records of a Swiss online shopping platform from July 2012 to November 2013; List of fraudulent accounts flagged by the retailer; Weekly advertisement speeding; Information of products available on the platform, etc.
Empirical methods: Vector autoregressive models; Survey		
Main results: - Based on the generalized impulse response function, unethical customer behavior could produce positive effects for retailers and peers. - Based on the survey, a majority of managers support keeping the unethical customers even through their behavior is unethical.		

Table 2: Overview of the second study in the thesis.

Study 2: *Fraudulent behavior and statistical fraud detection technique: A review*

Research questions	**Core contributions**	**Data basis**
What are the causes and consequences of unethical behavior? What techniques can we use to detect fraudulent behavior? How should the managers do to reduce the unethical behavior of their customers and employees?	Various articles from different disciplines have been reviewed to get a deeper understanding of unethical behavior. Highlight the importance of studies on ordinary unethical behavior. Various articles from different areas have been reviewed to get an idea of fraud detection techniques. Highlight the importance of network science in fraud detection.	Articles from the standard economics, psychology, behavior economics, and neuroscience on unethical behavior; Articles from computer science, physics, and network science on fraud detection techniques.

Main results:

- From the standard economics perceptive committing fraud does not have any internal cost.
- Psychology and behavior economics suggest that individuals do have internal costs while committing fraud. These costs emerge because the actions an individual undertakes, while engaging in cheating behavior, are inconsistent with the internal ethical standards and moral principles.
- The above-mentioned costs are different among individuals. Therefore, there exists no universal formula to predict precisely whether an individual commits fraud or not in a certain situation.
- Due to self-serving biases and bounded ethicality, people are sometimes not even aware of crossing ethical borders.
- Further development in neuroscience might shed light on understanding the underlying mechanisms behind these internal psychological costs.

Table 3: Overview of the third study in the thesis.

Study 3: *A Comparative Analysis of Community Detection Algorithms on Artificial Networks*		
Research questions Which is the most suited community detection algorithm in most circumstances based on observable properties of the network under consideration?	**Core contributions** Providing actual techniques to determine the most suited community detection algorithm. Proposing that the mixing parameter can be an easily measurable indicator of finding the ranges of reliability of the different algorithms. Showing the dependency with network size focusing on both the algorithm's predicting power and the effective computing time.	**Data basis** LFR (Lancichinetti, Fortunato & Radicchi) benchmark graphs.
Methods: Eight community detection algorithms available in the *"igraph"* package: Edge betweenness, Fastgreedy, Infomap, Label propagation, Leading eigenvector, Multilevel, Spinglass, and Walktrap.		
Main results: - For small networks, the community detection algorithms should be chosen based on their accuracies. - For large networks, algorithms should first be able to detect the organization of nodes in a reasonable time, and then have good accuracies. - By taking both accuracy and computing time into account, the "Multilevel" algorithm outperforms all the other algorithms on the set of benchmarks we have examined.		

References

[1] Babakus, E., Bettina, C. T., Mitchell, V. & Schlegelmilch, B. Reactions to unethical consumer behavior across six countries. *Journal of Consumer marketing* **21**, 254 – 263 (2004).

[2] Banerjee, Robin Who Cheats and How? In *Scams, fraud and the dark side of the corporate world, Chapter 2, Bank's Fraud Stories* (2015).

[3] Blondel, V. D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment* **2008**, P10008 (2008).

[4] Chen, Z., Hendrix, W. & Samatova, N. F. Community-based anomaly detection in evolutionary networks. *Journal of Intelligent Information Systems* **39**, 59–85 (2012).

[5] Clauset, A., Newman, M. E. & Moore, C. Finding community structure in very large networks. *Physical review E* **70**, 066111 (2004).

[6] Donaldson, T. & Dunfee, T. W. Toward a unified conception of business ethics: Integrative social contracts theory. *Academy of management review* **19**, 252–284 (1994).

[7] Ethics Alarms. The VW Scandal: Huge Consequences, Simple Ethics Lessons, Ominous Implications (2015). URL `https://ethicsalarms.com/2015/09/27/the-vw-scandal-huge-consequences-simple-ethics-lessons-ominous-implications/`.

[8] Girvan, M. & Newman, M. E. Community structure in social and biological networks. *Proceedings of the national academy of sciences* **99**, 7821–7826 (2002).

[9] Lancichinetti, A. & Fortunato, S. Benchmarks for testing community detection algorithms on directed and weighted graphs with overlapping communities. *Physical Review E* **80**, 016118 (2009).

[10] Lancichinetti, A., Fortunato, S. & Radicchi, F. Benchmark graphs for testing community detection algorithms. *Physical review E* **78**, 046110 (2008).

[11] Muncy, J. A., Vitell, S. J. Consumer ethics: An investigation of the ethical beliefs of the final consumer. *Journal of business Research* **24**, 297 – 311 (1992).

[12] Newman, M. E. Finding community structure in networks using the eigenvectors of matrices. *Physical review E* **74**, 036104 (2006).

[13] Pons, P. & Latapy, M. Computing communities in large networks using random walks. *J. Graph Algorithms Appl.* **10**, 191–218 (2006).

[14] Raghavan, U. N., Albert, R. & Kumara, S. Near linear time algorithm to detect community structures in large-scale networks. *Physical review E* **76**, 036106 (2007).

[15] Reichardt, J. & Bornholdt, S. Statistical mechanics of community detection. *Physical Review E* **74**, 016110 (2006).

[16] Rosvall, M. & Bergstrom, C. T. An information-theoretic framework for resolving community structure in complex networks. *Proceedings of the National Academy of Sciences* **104**, 7327–7331 (2007).

[17] Savage, D., Zhang, X., Yu, X., Chou, P.& Wang, Q. Anomaly detection in online social networks. *Social Networks* **39**, 62–70 (2014).

[18] UK Essays. Deshopping And Other Unethical Consumer Behaviours (2015). URL https://www.ukessays.com/essays/marketing/deshopping-and-other-unethical-consumer-behaviours-marketing-essay.php.

[19] Your Dictionairy. Examples of Unethical Behavior (2016). URL http://examples.yourdictionary.com/examples-of-unethical-behavior.html.

2 When Unethical Customers Have Beneficial Effects: A Theory of Retailer Response to Unethical Customer Behaviors[1]

Abstract

Conventional wisdom, supported by deontological ethics, suggests that retailers should extinguish unethical customer behavior when they discover it. We explore the boundaries of this widely-held view in two ways. First, we paint a more nuanced picture by proposing that when longer-term impact is considered, some unethical customer actions have beneficial consequences. We verify this core postulate by using a longitudinal dataset, covering seventy weeks and over 48,000 accounts, from a popular Swiss online retailer. Our results reveal that customers registering multiple accounts in violation of the retailer's policy comprise fewer than 11.5% of accounts, yet generate 27.5% of the retailer's revenue. Their participation leads to increased revenues for the retailer and increased engagement for other customers. Second, we theorize that when made aware, retailers will respond to unethical customer behavior in ways that accommodate this nuanced view. Supporting our prediction, a survey of retail managers finds that a vast majority (80.9%) are inclined to let unethical customers continue. However, ethical idealists are more likely to take a stand against unethical customers despite benefits accruing to the retailer.

key-words

Unethical customer behavior; marketing ethics; online shopping; loss prevention; consumer fraud; gamification.

[1] **Author Statement:** This work has been done together with René Algesheimer, and Utpal Dholakia. It's currently under the second round review at Journal of Retailing. We would like to thank the senior executives of the Swiss shopping site for supporting the research and for providing the data for the study. The financial support of the University Research Priority Programs (URPP) on social networks from the University of Zurich to Zhao and René and the financial support of the Jones Graduate School of Business, Rice University, to Utpal are gratefully acknowledged. Zhao Yang is the 1^{st} author, René Algesheimer is the 2^{nd} author, and Utpal Dholakia is the 3^{rd} and corresponding author.

2.1 Introduction

> "Rationality in economic ethics is bounded in three ways: by a finite human capacity to assess facts, by a limited capacity of ethical theory to capture moral truth, and by the plastic or artifactual nature of economic systems and practices."
>
> (Donaldson and Dunfee 1994, p. 258.)

Unethical customer behavior, defined as instances where consumers intentionally behave in a deceptive and dishonest ways that violate widely held moral principles, or disobey the retailer's rules or policies, takes many forms [11, 29, 32, 37, 54, 55]. Common instances include shoplifting, returning purchased items for a refund after use, accidentally or willfully damaging in-store merchandise, and providing false or misleading personal information such as Social Security number or telephone number [6, 24, 47]. Unethical behavior by some customers can produce significant ramifications for retailers and other consumers.

Given the potential severity of adverse consequences, unsurprisingly, much of the scholarly and popular attention has focused on negative aspects of unethical customer behavior and how to deal with it. Researchers have considered the antecedents and mechanisms through which customers decide to behave unethically, the range of negative consequences that occur, and ways of mitigating unethical behavior and its negative effects. This work yields recommendations for prevention or deterrence of unethical customer behavior through such means as impulse resistance, education, affecting social norms, and administering punishments (see Mazar and Ariely 2006, and Vitell 2003, for reviews [38, 54]).

In this paper, our central thesis is that such a perspective, although often practically useful, provides a narrow, and sometimes misleading, characterization of the consequences of customers' unethical behavior and considerations of how retailers should respond to them. We probe the boundaries of the widely-held view about unethical customer behavior in two ways. First, we develop a more nuanced and broader picture of unethical customer behavior by postulating that not all such behaviors are harmful. Some unethical actions can have beneficial consequences. Second, we develop a theoretical framework of retailer response that accom-

modates this nuanced view of unethical customer behavior and provides retailers with a wider assortment of response options.

Our theory makes the provocative and novel assertion that in some cases, an initial unethical customer action will have positive longer-term effects on the retailer and on other customers. Using a longitudinal dataset, covering seventy weeks, and over 48,000 accounts, from a popular Swiss online retailer, we verify this core postulate of our theory. In conducting the study, we extend investigations of unethical customer behavior beyond shorter-term decisions and actions usually examined with controlled lab experimentation [7, 33, 37, 48, 20] to a longitudinal field study of its consequences on a retailer's actual revenues and on its customers' actual purchase behaviors and activity over a period spanning more than a year. Our study is conducted in an online retailing context involving discounted purchases and customer-to-customer interactions around collecting and trading virtual cards. We utilize a list of accounts identified as behaving unethically (registering multiple accounts on the site by providing false information) through a robust multiple-stage verification method [8].

We also directly address the challenging question of how retailers should react when they encounter unethical customer behavior with predominantly longer-term positive effects. In such cases, retailers face a difficult trade-off: Should they stick to practice and prosecute or fire unethical customers even if doing so means hurting revenues and adversely affecting other customers' activities? Or should they ignore or even encourage such unethical behaviors? To the extent that the net effects on the retailer and other customers are positive, we argue that retailers will lean towards encouraging beneficial unethical behaviors.

We investigate this issue by conducting a survey of retail managers asking them how they would deal with this situation. The results reveal that consistent with our prediction, a majority of surveyed retail managers (80.9%) are inclined to let these unethical customers continue. However, those with greater ethical idealism are more likely to take a stand against unethical customers despite accruing benefits.

The rest of the paper is organized as follows. In the next section, we develop our theoretical framework, and describe its key postulates. Next, we describe our main study's setting

and explain how unethical consumer behavior is defined, detected, and tracked by the retailer. Our modeling approach is explained next followed by the results. After that, we present the results of a second study conducted with retail managers. We conclude with a discussion of the contributions and implications of our work.

2.2 A Theory of Retailer Response to Unethical Customer Behavior

In understanding how retailers conceive of, and deal with unethical customer behaviors, two opposing moral philosophy perspectives provide a useful starting point. The ethics literature distinguishes between deontological and teleological perspectives regarding what constitutes ethical behavior and how others should react to it [30, 27, 54]. A deontological perspective, exemplified by Kantian ethics [31], focuses solely on the inherent rightness or wrongness of an action and disregards its consequences. This focus arises from shining the light on the individual's motives for acting with the ultimate goal of behaving in a certain way for the right reasons. Contrarily, in assessing the ethicality of a particular behavior, a teleological perspective, with conceptual foundations in the Utilitarianism School developed by British philosophers Jeremy Bentham and David Hume [5], focuses on the cumulative positive and negative effects of the behavior's consequences and not on the nature of the behavior itself. It argues that a behavior can be considered as moral and should be encouraged as long as its total beneficial consequences outweigh its harmful consequences.

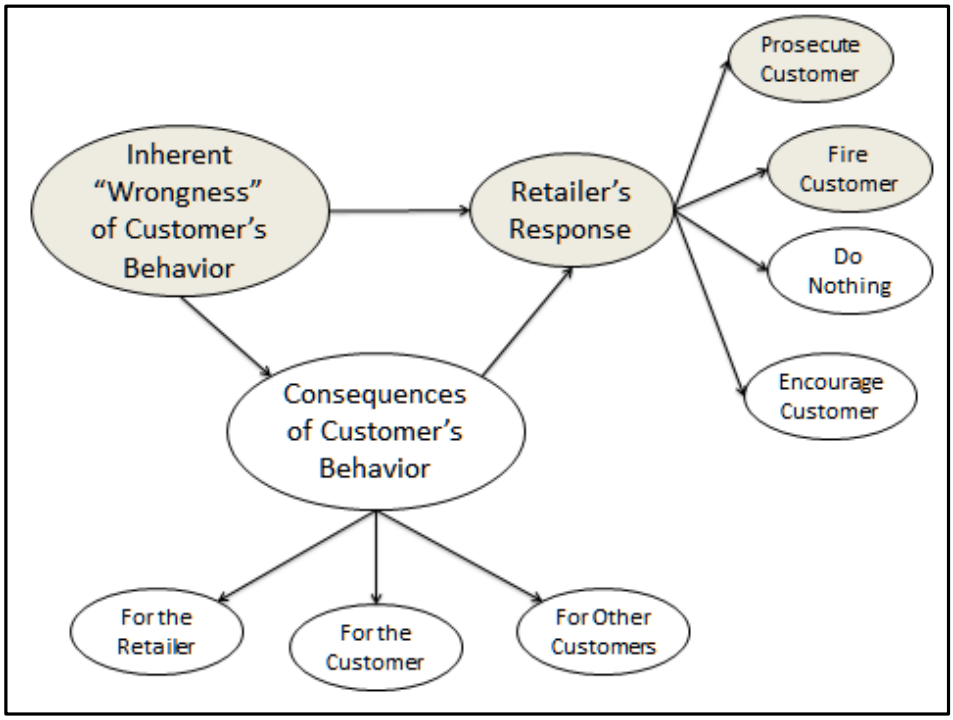
Interestingly, the conventional wisdom in the retailing literature as well as retailer practice itself endorses a deontological view. Retailers see all unethical customer behaviors as dysfunctional [2, 4, 47] and harmful to the bottom line, partly because they focus on the behavior itself which is more readily visible rather than the consequences of the behavior that are often harder to discern. Not surprisingly, this leads to the belief that any time unethically behaving customers are detected, retailers should use any and all means at their disposal to curtail the behavior. The more "wrong" the consumer's action in the sense of crossing ethical and legal

lines, the more strictly the retailer should deal with it [25, 26, 49]. To address the problem of shoplifting, for example, retailers commonly employ security personnel called "loss prevention officers." They not only prosecute shoplifters who are caught but also sue them for damages in civil court [4, 10]. This conventional view of dealing with unethical customers by prosecuting them is depicted in the top of Figure 1, marked in gray. It endorses the view that to the extent consumers behave unethically, such behaviors should be curbed by the retailer through legal and financial deterrence. Implicit in these actions is the idea that unethical customer behavior causes harm.

We advance the core postulate that, when deciding how to respond, the retailer should look beyond the unethical act itself, and consider longer-term consequences of the action, not only for itself but also for its other customers. In line with a teleological ethical perspective, and contemporary views of customer relationship management [36], we argue that customer behaviors that violate retailer policy and are marked as unethical can have a range of consequences, both negative and positive. What is more, these consequences stretch out over the customer's relationship with the retailer. In deciding how to respond to unethical behavior, it is important to distinguish between consequences for the retailer (e.g., effects on sales), consequences for the unethical customer (e.g., how the unethical action affects subsequent behavior), and consequences for other customers (e.g., how they are affected by the unethical customer's behavior). These different consequences are shown at the bottom of Figure 1.

Such a formulation of retailer response to unethical customer behavior requires two conditions to be met. First, the customer's unethical behavior should be innocuous enough to warrant a subdued or even a positive retailer response. When an unethical behavior is serious (e.g., assaulting a fellow customer in a store, stealing outright, etc.), the retailer must respond aggressively without teleological considerations. In other words, a range of retailer responses that support doing nothing or encouraging unethical customers under a teleological perspective presupposes that the customer's transgression was minor. The second condition is that the customer's unethical behavior should lead to measurable positive consequences for the retailer and other customers.

Figure 1: Theory of Retailer Response to Unethical Customer Behavior.



Psychological research provides evidence for both these conditions. Concerning the first condition, a growing body of literature over the past two decades provides evidence of "ordinary unethical behavior" [18], by showing that not only do significant numbers of individuals behave unethically in many everyday activities, but also that most such behaviors are small in both scale and seriousness [29, 38, 42, 54]. The reasons lie in the nuanced decision calculus that instigates unethical behavior. Individuals balance the need to profit from behaving unethically with the need to maintain a positive self-view. Consequently, the extent of their unethical behavior is limited to minor infractions – actions seen as relatively innocuous (that is, not causing substantial harm to anyone), and where such actions can be easily justified or rationalized to oneself [32, 37] or where they can "strategically forget" the ethical principles involved [50]. Based on this research, we may surmise that retailers are likely to frequently encounter customers who conduct minor ethical transgressions. But can such actions have beneficial effects?

There is emerging psychological research that has begun to study positive effects of one's unethical behavior on others. Erat and Gneezy introduce the concept of a "pareto white lie" in

which both the liar and others benefit [16]. They offer the example of a physician knowingly lying and giving a placebo to patients, knowing fully well that such a medicine will have no pharmacological effect on the patient, but may confer psychological benefits leading to positive outcomes. The medical literature argues that such lying by medical practitioners constitutes "benevolent deception", and is morally defensible to the extent that it improves the patient's care and increases likelihood of positive patient outcomes [28]. However, no studies, to our knowledge, have considered potential benefits of unethical customer behaviors in retailing or marketing contexts.

We theorize that in relational settings such as typical retailing environments, an initial unethical action by the customer may open the door for him or her to behave in legitimate ways on an ongoing basis. Furthermore, such activities may add substantial value to both the retailer and other customers, outweighing the harm of the person's initial unethical action. When a range of possible outcomes from the negative to the positive is possible, the retailer must respond according to the consequences. Thus, as shown in the figure, once the consequences are measured and understood, the retailer will choose to prosecute the customer, fire him or her, do nothing, or even encourage the customer depending on the positivity and negativity of the consequences [15].

Our theory of retailer response to unethical customer behavior makes two empirically testable postulates that we investigate in this paper. The first postulate, that forms the focus of our main study, is that customers behaving unethically at the outset can contribute positively to the retailer and to other customers when their behaviors are considered and evaluated over the longer-term. The second postulate is that when retailers understand the potential positive longer-term ramifications of customers' unethical behavior, they will be likely to adopt a more measured response to this behavior than is predicted by a strictly deontological ethical view that are implicitly endorsed by many retailers.

2.3 Study Setting

This study is conducted in Switzerland and employs longitudinal data from a Swiss online retailer that was founded in late 2011. Focused primarily on retail shopping for generating revenue, the site provides its customers with an engaging shopping experience by using social gaming and price promotions [57]. These features serve dual purposes of distinguishing the site in a crowded retailing environment, and encouraging customers to visit regularly. Customers of this site actively collect and trade virtual cards associated with each offer. In return, they receive discounts corresponding to the number of cards they have collected at the time of purchase. Since its founding, the firm's customer base has grown at an annual rate of 114%. By the end of 2014, it had more than 100,000 active customers.

The Retailer's Business Model

The site sells a variety of goods and services. Products offered over the course of this study included the Samsung Galaxy, the Apple iPad, Sony Xperia, various branded clothes and handbags, and services such as pre-paid salon and spa services, restaurant meals, and trips. When an offer is first listed on the site, a set of ten virtual cards (numbered from 1 to 10) is generated for the offer. Each card corresponds to a discount voucher that offers a ten percent discount off the item's listed price. Customers have the option of either purchasing the item at the full listed price or collecting cards associated with the offer to bring down the price they will have to pay. As the customer collects more cards associated with an offer, the discount received off the list price increases additively. Thus, if the customer is able to collect all ten cards, they receive the listed item for free. The customer is free to redeem all cards they have collected at any time during the offer period and receive a discount corresponding to the number of collected cards. Most offers last 4-6 weeks, giving customers with time to collect and trade cards with each other.

The Role of Virtual Trading Cards

Customers can obtain virtual trading cards in a number of ways. First, when new users join the site, after registering and logging for the very first time, they receive three randomly chosen cards. Second, to encourage regular visits, each account receives two free randomly chosen cards on a daily basis for simply logging in. Note that the customer must log in at least once during the day to receive these free daily cards. On days that they do not log in, they don't receive any cards. If they log in more than once on any given day, they still receive only two cards. Third, to augment these sources, customers have the ability to purchase cards from the site at the price of 2 Swiss francs (CHF) per card. These cards are generated randomly; the customer cannot specify or order a particular card for purchase. It is important to note here that there is no difference in the generation process for cards that are given for free and those that are paid for by customers. All cards are generated using the same random process. However, for an offer, each of the ten cards is generated with a different probability and has a different cap, so that some cards are quite common while others are very rare. The total number of released cards also varies from one offer to the next depending on the number of items that the site wants to sell and the stipulations of the manufacturer. Because of these characteristics, customers may find that for one particular offer, some cards are easy to get whilst others are extremely difficult.

Fourth and importantly, customers can trade cards they possess with other site members. This gamification feature is the site's key differentiator when compared to other shopping and discount voucher sites. Specifically, collecting and trading cards with each other to earn greater discounts on a particular offer not only introduces a social aspect to the customer's experience but it also involves the explicit use of gamification to produce excitement. Both enhance customers' engagement with the site and with each other. To encourage card trading, a search engine is prominently displayed on the site's main page. Using this engine, customers can find other members with a specific card they are looking for. They can send an offer for the desired card through the site's messaging system and negotiate once contact is established and if there is mutual interest in a trade. If a customer desires one particular card intensely, he or she can

offer multiple cards in return. Finally, once a relationship is established between customers, they can trade cards with each other directly without using the site's search engine. There is no limit to how many cards the transaction parties can submit to a trade.

Definition of a Fraudulent Account

By lowering the price that the customer has to pay for the offered item by 10%, each trading card has a significant, measurable financial value. Specifically, the financial value of a card equals a tenth of the offer's value. However, due to varying availability of cards, scarcity drives the perceived financial value for card collectors. To provide equal chances to its customers of receiving all cards, and to comply with Swiss lottery law, the retailer requires each customer to register when joining the site. It stipulates that each person can open only one account, and the account must be associated with only that person. These stipulations are clearly described in the retailer's published terms and conditions since its inception and are also explained to customers in simple language in email communications.

In this research, we study the effects of customers who violate this policy, i.e. behaving unethically. In the analysis that follows, we define fraudulent users as customers who misrepresent themselves to take advantage of free cards offered to each user (upon registration and when logging in daily), thereby winning a stronger bargaining position in the trading card game. They do this by opening multiple accounts using different mobile or VOIP telephone numbers (e.g., such as those belonging to their family or friends) to confirm each of their accounts. Such a misrepresentation is considered as fraudulent behavior not only because it violates the retailer's terms and conditions, but because it could precipitate a potential monetary loss for other legitimate customers by decreasing their likelihood of winning a particular trading card offered [8].

From the customer's viewpoint, there are at least four advantages to opening and maintaining multiple accounts. First, the more accounts one has, the more daily free cards one receives, increasing the chances of winning multiple cards associated with the same offer and earning a higher cumulative discount. Additionally, because they have more free cards and continue to

get them at a faster rate, individuals with multiple accounts have the flexibility to trade extravagantly, offering multiple cards in exchange for a particular rare card. Second, having multiple accounts also increases the person's chances of receiving rare cards having greater value for trades. Third, the more cards a customer collects, the easier it is to recognize a card's scarcity value, leading to a position of strength in negotiating trades. Fourth, customers with multiple accounts can trade and exchange cards between their own accounts. For these reasons, the retailer sees multiple accounts as unfair to its legitimate customers who maintain a single account.

Identifying Fraudulent Accounts

To identify fraudulent customers, the retailer employs a multi-step procedure. Since its inception in late 2011, it has used Google Analytics. As part of gathering behavioral data from shoppers, the site serves cookies to customers' computers and mobile devices, which are then used to detect fraudulent accounts (multiple accounts accessed from the same device; see "Supplementary Information of Chapter 2" for details regarding identifying fraudulent accounts). Accounts are observed for their fraud potential based on activity. They are flagged if two (or more) accounts with similar profiles routinely log on from the same IP address within a few minutes of each other.

Once accounts are flagged, the retailer contacts the account owners by telephone to determine whether it was the same person associated with the different accounts (to rule out legitimate same-IP users such as spouses or partners, children, room-mates, etc.). With this manual process, the retailer is able to continually modify and improve the accuracy of its fraud potential algorithm. If this process fails to clear the customer, at this stage, the account is flagged as fraudulent and the retailer begins a protocol for having the customer delete multiple accounts. First, it will contact accounts identified as fraudulent by email, and next, it will do so through in-platform messages to issue warnings. If these two communications also fail, in the third stage of the protocol, the retailer confronts the account owner with their suspicion, and invites them to respond within a limited period of time (e.g., four weeks) before terminating their accounts.

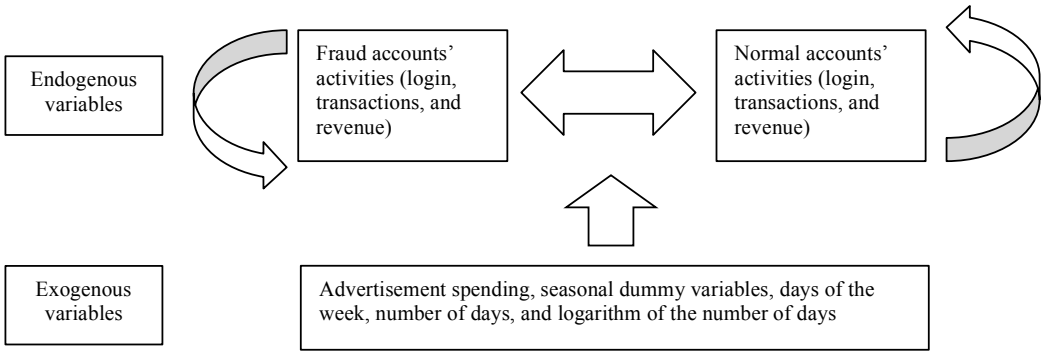
2.4 Data and Modeling Approach

The retailer provided us with longitudinal data for 48,782 unique accounts that engaged in 3,083,936 logins, 2,022,666 total transactions, and 316,564 successful transactions, along with the retailer's weekly advertisement spending during the time period from 4 July 2012 to 22 November 2013. There are a total of 19,416 customer purchases during this time that included both buying trading cards (at a price of 2 CHF per card) and purchasing products and services (with or without accompanied card redemptions). During this period, the retailer presented a total of 544 different offers on its site, with product list prices ranging from 11 CHF to 6,882 CHF (M = 338.2 CHF). Due to site maintenance and hardware upgrade issues there was a period of 18 days during which the site was shut down during January 2013. Important for our purposes, among these 48,782 accounts, we also received a list of 5,576 unique accounts flagged as fraudulent by the retailer (accounts that failed to clear the manual verification process).

We use vector autoregressive models to study the effects of fraudulent accounts on normal accounts. The impact of fraudulent customer behavior on the site on other user's behavior using is modeled the following approach. First, we test for potential endogeneity among variables related to fraudulent and normal accounts. Next, we specify a vector autoregressive model (VAR) accounting for endogeneity and the fully interacting dynamic system of endogenous variables. After that, we estimate the generalized impulse-response function to study the impact of fraudulent accounts on normal ones. Finally, we compare our VAR model to a number of alternative models. Figure 2 describes our modeling framework graphically, with the arrows in the figure indicating the links between endogenous and exogenous variables.

To measure *customer activity*, we employ the number of logins, the number of proposed transactions, and the number of successful transactions. In the number of proposed transactions, we have only considered the transaction requests proposed by the accounts, and excluded those received from others. This is due to the fact that customers cannot prohibit trading requests from others; thus received transaction requests do not reflect any activity on the customer's part. However, in the number of successful transactions, we have counted both transactions proposed

Figure 2: The VAR Modeling Framework.



and received by the accounts because for successful transactions, customers are actively involved on both sides of the transaction. Revenue is measured as the *amount of money spent on the site*. The retailer's advertisement spending, the seasonal dummy variables, days of the week, and time trends were included as *controls*. We transformed the weekly advertisement spending into daily level by assuming that the advertisement spending is equally distributed throughout the week. Additionally, lagged effects are included in the model (See "Supplementary Information of Chapter 2" for details regarding the lag-order selection). As we are investigating the interactions of groups of fraudulent accounts and normal accounts, we aggregate all individual accounts into the fraudulent accounts group and the normal accounts group. The analysis is conducted on a daily level.

It is important to test for the presence of endogeneity between customer activity and the site's revenue performance. As Figure 2 indicates, we anticipate that the activity of fraudulent accounts will affect the activity of normal accounts and the change in normal accounts' activity will correspondingly lead to changes in fraudulent accounts' activity and will therefore indirectly impact the change in normal accounts' activity. We anticipate similar patterns of causality between the revenue of fraudulent accounts and the revenue of normal accounts, and between the activity and the revenue. The links represented in Figure 2 can be tested by investigating which variables Granger cause other variables [21, 23].

In essence, the Granger causality test examines whether one variable temporally causes a second variable after accounting for the history of the second variable. In the absence of controlled experimentation, such a "temporal causality" is the closest proxy for causality that can be obtained from time-series data. It is important to point out that a wrong choice for the lag-order in the test may erroneously conclude the absence of Granger causality [22]. Because we are applying these tests to investigate the need for modeling a full dynamic system, we are not interested in whether variable X causes variable Y at a specific lag, but in whether we can exclude that X Granger causes Y at any lag [53]. Therefore, we conduct the Granger-causality tests on each pair of variables for all the lags up to 30 with the null hypothesis that the variables do not Granger cause other variables included in the analysis. The minimum p-values for the lag that has the highest significance for Granger causality are reported in Table 4. In summary, the results from the Granger-causality tests indicate the need to consider the full dynamic system, as in a VAR model, and to account for the indirect effects of different actions.

Table 4: Results of the Granger-causality tests.

Results of the Granger Causality Tests (Minimum p-Value Across 30 Lags)								
Dependent Variable Is Granger-Caused by	n_PT_t	n_ST_t	n_L_t	n_T_t	f_PT_t	f_ST_t	f_L_t	f_T_t
n_PT_t	—	.000	.000	.000	.066	.067	.101	.000
n_ST_t	.000	—	.000	.002	.002	.002	.000	.000
n_L_t	.000	.000	—	.000	.001	.000	.000	.000
n_T_t	.000	.000	.000	—	.000	.000	.000	.000
f_PT_t	.000	.000	.000	.000	—	.000	.000	.000
f_ST_t	.000	.000	.000	.000	.017	—	.000	.000
f_L_t	.000	.000	.000	.000	.000	.000	—	.000
f_T_t	.000	.000	.000	.000	.000	.001	.000	—

Notes: n_PT_t is the number of transactions proposed by normal accounts at day t; n_ST_t is the number of successful transactions made by normal accounts at day t; n_L_t is the number of logins of normal accounts at day t; n_T_t is the amount of revenue of normal accounts at day t; f_PT_t is the number of proposed transactions of fraudulent accounts at day t, f_ST_t is the number of successful transactions of fraudulent accounts at day t; f_L_t is the number of logins of fraudulent accounts at day t; f_T_t is the amount of revenue of fraudulent accounts at day t.

Next, to determine whether each of the variables in our dataset is stable (i.e. whether it fluctuates temporarily around a fixed mean or trend) versus evolving (i.e. whether it can deviate permanently from the previous level), we apply the Augmented Dickey-Fuller test [14] and the Phillips-Perron test [46] to the time series. The null hypothesis of both tests is that the variable contains a unit root (the variable is not stationary), and the alternative is that the variable is generated by a stationary process. Tables 5 and 14 provide the results of these tests respectively, including the test statistics and the Mackinnon approximate p-value for the test statistics. These results reject the null hypothesis at the common significance levels, confirming that all variables appear stationary after controlling for trend and lagged difference. Thus we can perform our model estimations with the variables in levels.

Table 5: Augmented Dickey-Fuller unit root test results.

Augmented Dickey-Fuller test (trend, lags(2)) (The p-value is in the last column)					
Variables	**Test statistics**	**1% Critical value**	**5% Critical value**	**10% Critical value**	**Mackinnon approximate p-value for Z(t)**
n_PT_t	-5.835	-3.981	-3.421	-3.130	.0000***
n_ST_t	-5.92	-3.981	-3.421	-3.130	.0000***
n_L_t	-5.776	-3.981	-3.421	-3.130	.0000***
n_T_t	-10.176	-3.981	-3.421	-3.130	.0000***
f_PT_t	-6.103	-3.981	-3.421	-3.130	.0000***
f_ST_t	-5.61	-3.981	-3.421	-3.130	.0000***
f_L_t	-4.696	-3.981	-3.421	-3.130	.0007***
f_T_t	-9.098	-3.981	-3.421	-3.130	.0000***

Significance codes: 0 "***", 0.001 "**", 0.01 "*", 0.05 " "

Encouraging results of the tests for both endogeneity and unit root allow us to proceed to specify and formalize the VAR model in Figure 3 to study the relationships between variables of fraudulent accounts and normal accounts (see "Supplementary Information of Chapter 2" for details regarding the stability test of the VAR model):

Where n_PT_t is the number of transactions proposed by normal accounts at day t; n_ST_t is the number of successful transactions made by normal accounts at day t; n_L_t is the number of logins of normal accounts at day t; n_T_t is the amount of revenue of normal accounts at

Table 6: Phillips-Perron unit root test results.

Phillips-Perron test (trend, lags(2)) (The p-value is in the last column)					
Variables	**Test statistics**	**1% Critical value**	**5% Critical value**	**10% Critical value**	**Mackinnon approximate p-value for Z(t)**
n_PT_t	-7.476	-3.981	-3.421	-3.130	.0000***
n_ST_t	-6.98	-3.981	-3.421	-3.130	.0000***
n_L_t	-6.956	-3.981	-3.421	-3.130	.0000***
n_T_t	-18.59	-3.981	-3.421	-3.130	.0000***
f_PT_t	-7.743	-3.981	-3.421	-3.130	.0000***
f_ST_t	-7.134	-3.981	-3.421	-3.130	.0000***
f_L_t	-5.176	-3.981	-3.421	-3.130	.0001***
f_T_t	-18.026	-3.981	-3.421	-3.130	.0000***

Significance codes: 0 "***", 0.001 "**", 0.01 "*", 0.05 " "

Figure 3: Formula of the VAR Model.

$$\begin{bmatrix} f\_L_t \\ f\_PT_t \\ f\_ST_t \\ f\_T_t \\ n\_L_t \\ n\_PT_t \\ n\_ST_t \\ n\_T_t \end{bmatrix} = \sum_{i=1}^{lag} \begin{bmatrix} \alpha^i_{1,1} & \cdots & \alpha^i_{1,8} \\ \vdots & \ddots & \vdots \\ \alpha^i_{8,1} & \cdots & \alpha^i_{8,8} \end{bmatrix} \begin{bmatrix} f\_L_{t-i} \\ f\_PT_{t-i} \\ f\_ST_{t-i} \\ f\_T_{t-i} \\ n\_L_{t-i} \\ n\_PT_{t-i} \\ n\_ST_{t-i} \\ n\_T_{t-i} \end{bmatrix} + \begin{bmatrix} \gamma^i_{1,1} & \cdots & \gamma^i_{1,12} \\ \vdots & \ddots & \vdots \\ \gamma^i_{8,1} & \cdots & \gamma^i_{8,12} \end{bmatrix} \begin{bmatrix} adspending_t \\ season1_t \\ season2_t \\ season3_t \\ dofWk1_t \\ dofWk2_t \\ dofWk3_t \\ dofWk4_t \\ dofWk5_t \\ dofWk6_t \\ \#days_t \\ \#log(days)_t \end{bmatrix} + \begin{bmatrix} \varepsilon_{f\_L,t} \\ \varepsilon_{f\_PT,t} \\ \varepsilon_{f\_ST,t} \\ \varepsilon_{f\_T,t} \\ \varepsilon_{n\_L,t} \\ \varepsilon_{n\_PT,t} \\ \varepsilon_{n\_ST,t} \\ \varepsilon_{n\_T,t} \end{bmatrix}$$

day t; f_PT_t is the number of proposed transactions of fraudulent accounts at day t, f_ST_t is the number of successful transactions of fraudulent accounts at day t; f_L_t is the number of logins of fraudulent accounts at day t; f_T_t is the amount of revenue of fraudulent accounts at day t; $adspending_t$ is the advertising spending of the retailer at day t; $season1_t$ to $season3_t$ are seasonal dummies; $dofWk1_t$ to $dofWk6_t$ are day of the week dummies; $\#days_t$ and $\#log(days)_t$ are time trends.

In this model, the revenue, the login, the number of proposed transactions, and successful transactions made by fraudulent accounts and normal accounts are endogenous they are explained by their own past and the past of the other endogenous variables [13]. The vector of exogenous variables includes advertising spending, seasonal dummy variables, day of week effect dummy variables, and time trend variables. The off-diagonal terms of the matrices $\alpha^i_{j,k}$ $(j \neq k)$ estimate the direct and cross-over effects among all endogenous variables, and diagonal elements $(j = k)$ estimate auto-regressive effects. The terms of the matrices $\gamma^i_{j,l}$ represent the impact of exogenous variables controlling for ad spending, seasonality, day of the week, and time trends.

2.5 Results

We begin by first reporting straightforward model-free evidence regarding the importance of fraudulent accounts to the retailer. After that, we provide estimates of the vector autoregressive model.

Summary Statistics

The descriptive statistics for our sample are provided in Table 7. By comparing the totals and averages in the table, we can conclude that the fraudulent accounts are among the site's most active accounts. 11.43% of accounts are fraudulent yet they contribute 27.66% of the site's revenues. Customers with fraudulent accounts appear to be more willing to login, to propose trade transactions, and to spend money on the site. Moreover, they are more successful in con-

summating trade transactions, as indicated by the higher ratio for successful transactions than proposed transactions in Panel B of table 7. They log in 5 times as often as normal accounts, propose 3.6 times as many transaction requests, achieve 6.1 times as many successful transactions, and spend 3 times as much as normal accounts.

Table 7: Statistics of normal accounts and fraudulent accounts in the final sample.

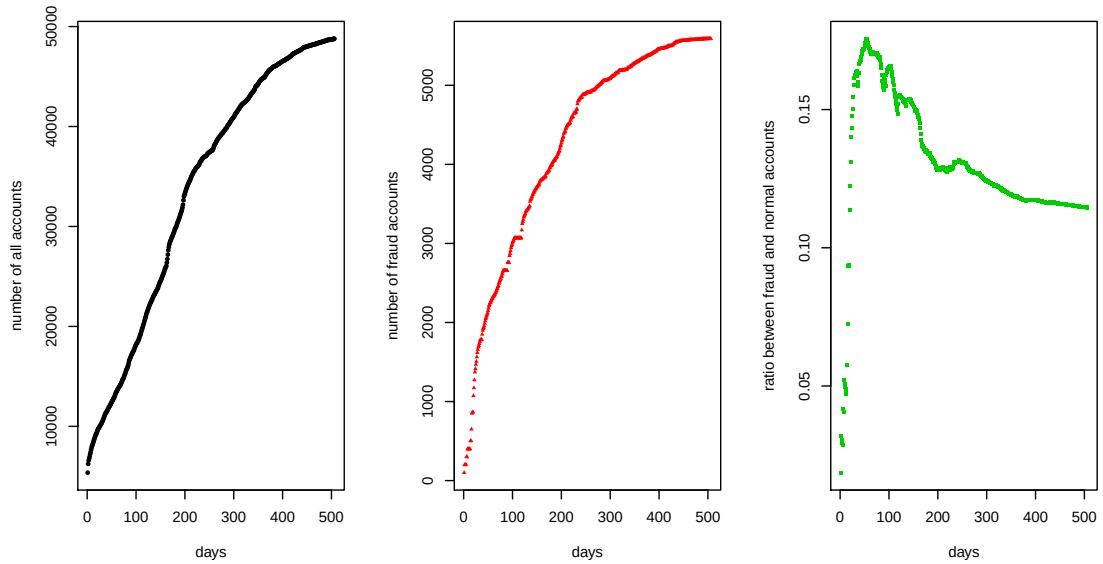
Panel A: Total Numbers			
Variables	**Normal accounts**	**Fraudulent accounts**	**Total accounts**
Number (% of sample)	43,206 (88.57%)	5,576 **(11.43%)**	48,782
Number of logins	1,877,782 (60.89%)	1,206,154 (39.11%)	3,083,936
Number of proposed transactions	1,382,994 (68.37%)	639,672 (31.63%)	$2,022,666^a$
Number of successful transactions	353,454 (55.83%)	279,674 (44.17%)	$316,564^b$
Spending on the site (CHF)	579,284.4 (72.34%)	221,476.35 **(27.66%)**	800,760.75
Panel B: Average Numbers			
Variables	**Normal accounts**	**Fraudulent accounts**	**Fraud Accts/ Normal Accts**
Average number of logins per account	43.46	216.31	4.98
Average number of proposed transactions per account	32.01	114.72	3.58
Average number of successful transactions per account	8.18	50.16	6.13
Average spending on the site per account (CHF)	13.41	39.72	2.96

[a]In the number of proposed transactions, only transaction requests proposed by the accounts are considered. Customers cannot prohibit trading requests from others, but these requests do not constitute the customer's own activity.
[b]In the number of successful transactions, both transactions proposed and received by the accounts are included because in a successful transaction, the customer is actively involved on both sides of the transaction.

Figure 4 shows the cumulative number of all accounts, fraudulent accounts, and their ratio over time. In the panels of the figure, once an account was flagged as fraudulent, it was considered as fraudulent from then on. From the right panel showing the ratio of cumulative

fraudulent accounts to all accounts, we can surmise that the retailer was able to detect fraudulent accounts quickly early on in its existence. As it employed protocols of emailing and then warning fraudulent customers, the rate of new fraudulent accounts was brought under control. The ratio stabilized and then declined as time passed.

Figure 4: Development of Number of Accounts Over Time.

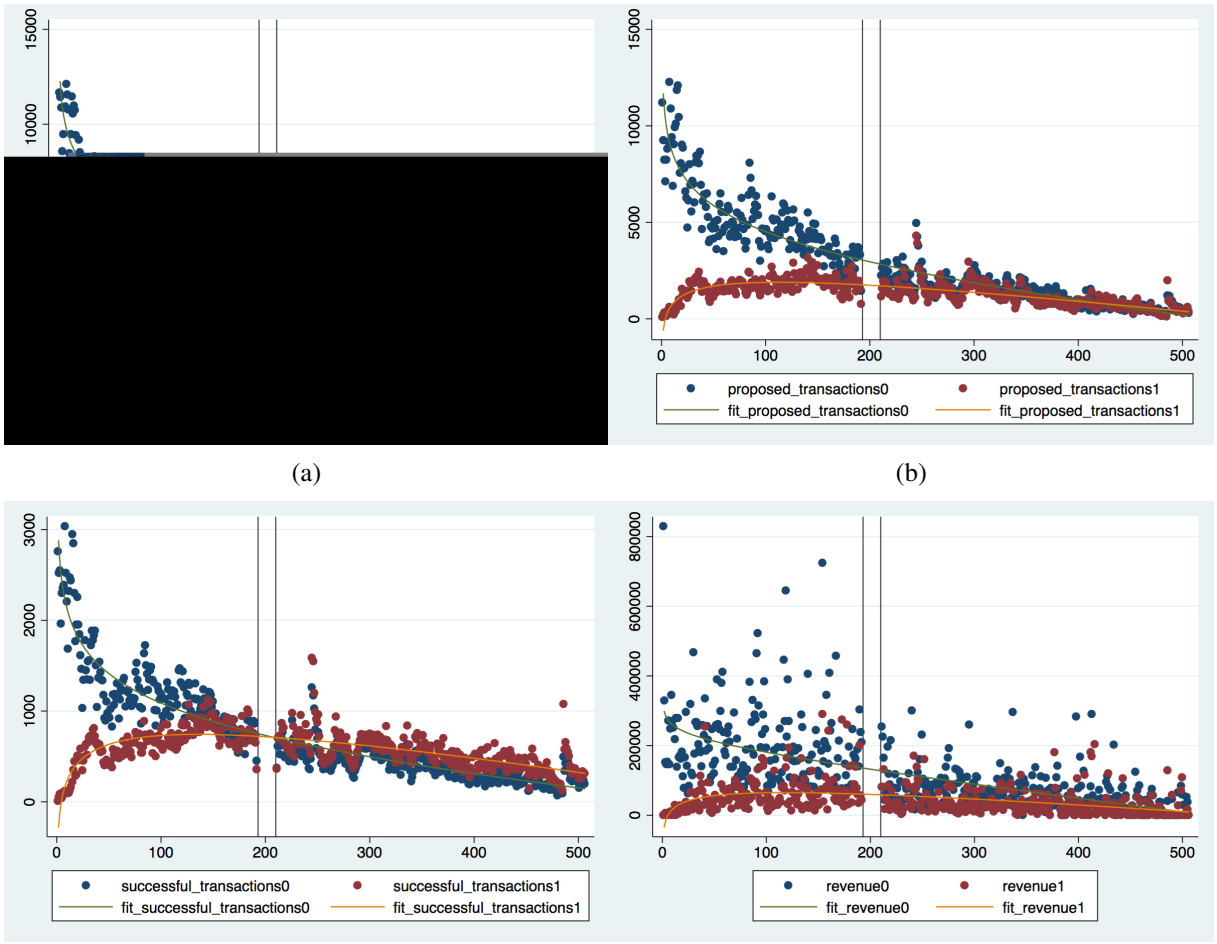


Results of the VAR Model: generalized impulse response function

In order to examine and understand the time trends in our data, we plot the number of logins, proposed transactions, successful transactions, and the retailer's revenue for the fraudulent and normal accounts over time. After testing a number of different functions, the time trends could be best fitted with t plus $log(t)$, where t is the number of days. Figure 5 provides the raw data and the fitted curves.

In the figure, fraudulent accounts are represented as red nodes (type = 1) and normal accounts are represented as blue ones (type = 0). The green curve is the fitting of normal accounts and the orange one is the fitting of fraudulent accounts. Both curves are fitted based on the function $y = t + log(t)$. The two black vertical lines represent the 18-day period during which

Figure 5: Levels of Variables For Fraudulent and Normal Accounts Over Time: (a) Number of Logins, (b) Number of Proposed Transactions, (c) Number of Successful Transactions, and (d) Revenue.



(a) (b)

(c) (d)

the site was down for service maintenance. Since there is no distinct effect of the maintenance on customer login activity, transaction activities, and purchasing behavior, we do not consider it further.

As it is difficult to interpret the coefficients estimated by the VAR model directly [51], we employ impulse response functions (IRFs) to simulate the impact of a change (over its baseline) in one variable over time on the full dynamic system, thus representing the net result of all modeled actions and reactions. The original IRFs are the "orthogonalized" impulse responses, where the underlying shocks to the VAR model are orthogonalized using the Cholesky decom-

position or forecast error variance decompositions before impulse responses. However, this approach is sensitive to the causal ordering of variables in the VAR model [35]. To solve this issue, we apply the generalized impulse response function which does not ask for a causal ordering among the endogenous variables but instead uses information available in the residual variance-covariance matrix of the VAR model [13, 45, 34].

Figure 6 displays the results of the generalized impulse response functions for the impact of an unexpected change in fraudulent accounts' login activity, the number of transactions proposed by fraudulent accounts, the number of successful transactions made by fraudulent accounts and the revenue resulting from fraudulent accounts on login activity of a normal account. We transform the IRFs into elasticity to show the impact of the positive shock (Trusov, Bucklin, and Pauwels 2009) as follows: First, the IRF analysis yields to the change in number of the variable, ΔY, in response to a positive one-standard-deviation shock to another variable, X. Second, we calculate the standard deviation for X (denoted as σ_x), the mean values for Y (denoted as $\bar{Y}$), and the mean values for X (denoted as $\bar{X}$). Finally we get the arc elasticity as shown in Equation 1:

$$\eta_{arc} = \frac{\Delta Y}{\sigma_x} \times \frac{\bar{X}}{\bar{Y}} \tag{1}$$

Based on the results, we can conclude that: (1) An unexpected positive shock on fraudulent accounts' login times will yield an increase of 159 logins on normal accounts' login activity. The corresponding elasticity is 0.1291, i.e. the normal accounts' login activity will increase by 12.91% in response to a positive one-standard-deviation shock to fraudulent accounts' login activity. This positive effect will decrease over time and disappear after 5 days; (2) An unexpected positive shock on fraudulent accounts' number of proposed transactions will yield an increase of 119 units on normal accounts' login activity. The corresponding elasticity is 0.0653 (6.53%). This positive effect will dissipate to zero after 4 days; (3) An unexpected positive shock on fraudulent accounts' number of successful transactions will yield an increase of 119

logins on normal accounts' login activity. The corresponding elasticity is 0.0865 (8.65%). This positive effect will last for 4 days before it becomes insignificant; and (4) an unexpected positive shock on fraudulent accounts' amount of revenue will not yield any significant response on normal accounts' login activity. Similarly, Figures 7, 8, and 9 show the same four impulses on the number of proposed transactions of normal accounts, the number of successful transaction made by normal accounts, and the revenue of normal accounts, respectively.

Suppose further that the login and transaction activities of the platform consist of activities of fraudulent accounts and normal accounts, and the retailer's revenue equals to the sum of the revenue from fraudulent accounts and normal accounts, we can also calculate the corresponding retailer elasticity as shown in Equation 2:

$$\eta_{arc} = \frac{\Delta X + \Delta Y}{\sigma_x} \times \frac{\bar{X}}{\bar{X} + \bar{Y}} \tag{2}$$

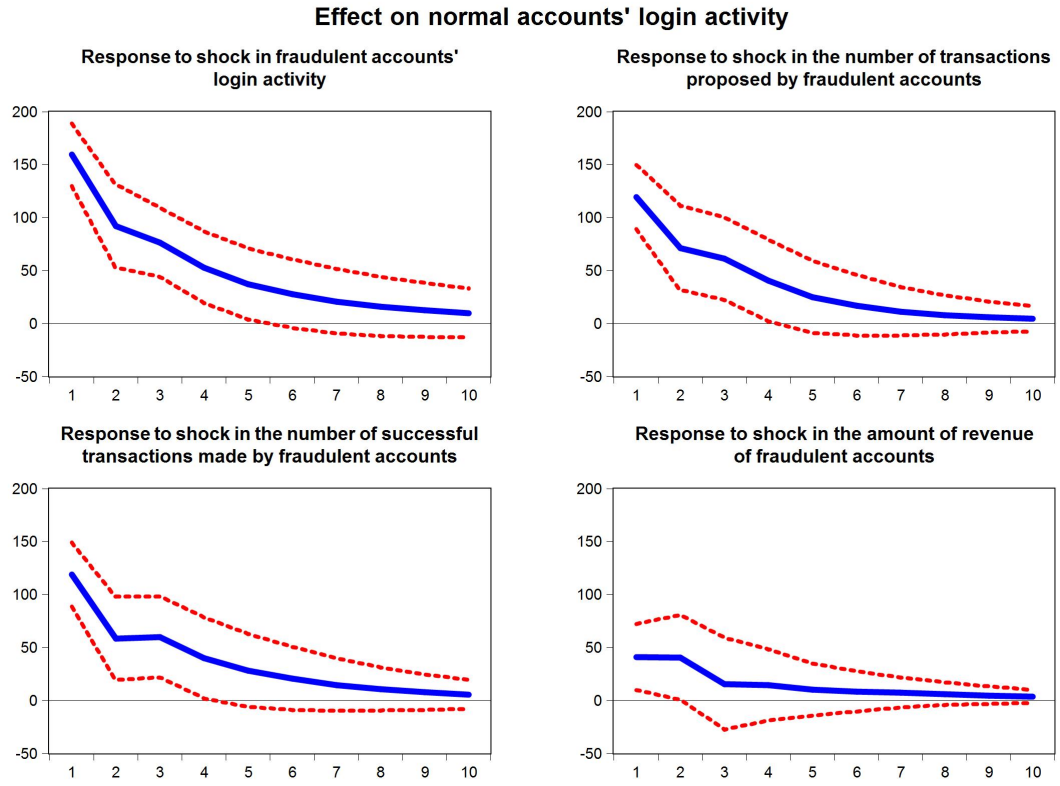
Table 8 summarizes the retailer's elasticity in response to the unexpected positive 1 standard deviation shock on fraudulent accounts and normal ones, separately. We can conclude that: (1) with respect to platform activities (number of logins, number of proposed transactions, and number of successful transactions), fraudulent accounts consistently contribute more than normal accounts, and (2) with respect to retailer revenue, fraudulent accounts make a less, but still significant contribution compared to the normal ones.

Results of the VAR Model: Restricted impulse response functions

As the generalized impulse response functions represent the net effect of the full chain of events set in motion by the shock, it is of interest to specify its direct effects versus the indirect ones. To do the effect separation, we adapt the idea of "conceptual experiments," in which one adds restrictions that only allow some variables to react, while keeping other variables at their baseline level [43, 44].

To obtain the long term direct effect of fraudulent accounts on normal ones, we estimate

Figure 6: The plots of the generalized impulse response functions for the impact of an unexpected change in fraudulent accounts' variables on normal accounts' login activity.



In the figure, the blue curve shows the impact and the red dotted curves represent the one standard deviation confidence interval for the impact. The x-axis is the number of days. The y-axis is the number of login times.

An unexpected positive shock on fraudulent accounts' login times will yield an increase of 159 login times on normal accounts' login activity. The corresponding elasticity is 0.1291. This positive effect will decrease over time and disappear after 5 days.

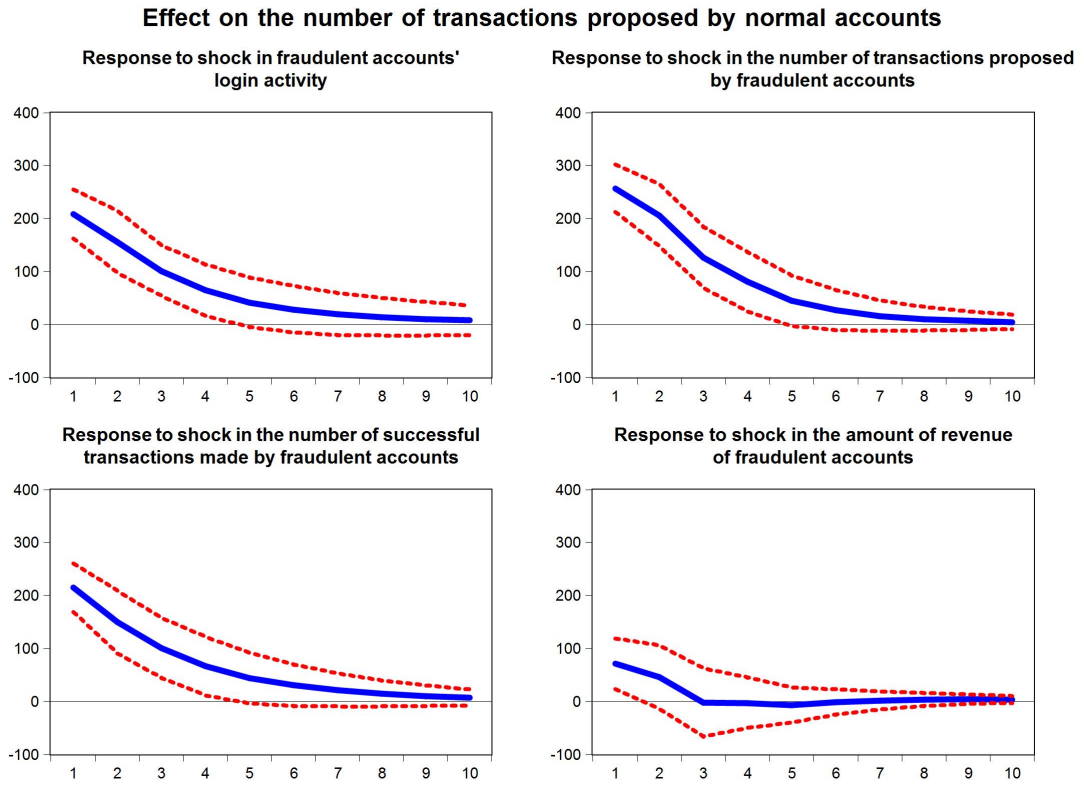
An unexpected positive shock on fraudulent accounts' number of proposed transactions will yield an increase of 119 login times on normal accounts' login activity. The corresponding elasticity is 0.0653. This positive effect will go to zero after 4 days.

An unexpected positive shock on fraudulent accounts' number of successful transactions will yield an increase of 119 login times on normal accounts' login activity. The corresponding elasticity is 0.0865. This positive effect will last for 4 days before it becomes insignificant.

An unexpected positive shock on fraudulent accounts' amount of revenue won't yield any significant response on normal accounts' login activity.

separate impulse response functions by restricting variables to remain unaffected by the corresponding shocks on fraudulent accounts. As the restricted impulse response functions and the generalized impulse response functions are based on the same estimated coefficients from the

Figure 7: The plots of the generalized impulse response functions for the impact of an unexpected change in fraudulent accounts' variables on the number of transactions proposed by normal accounts.

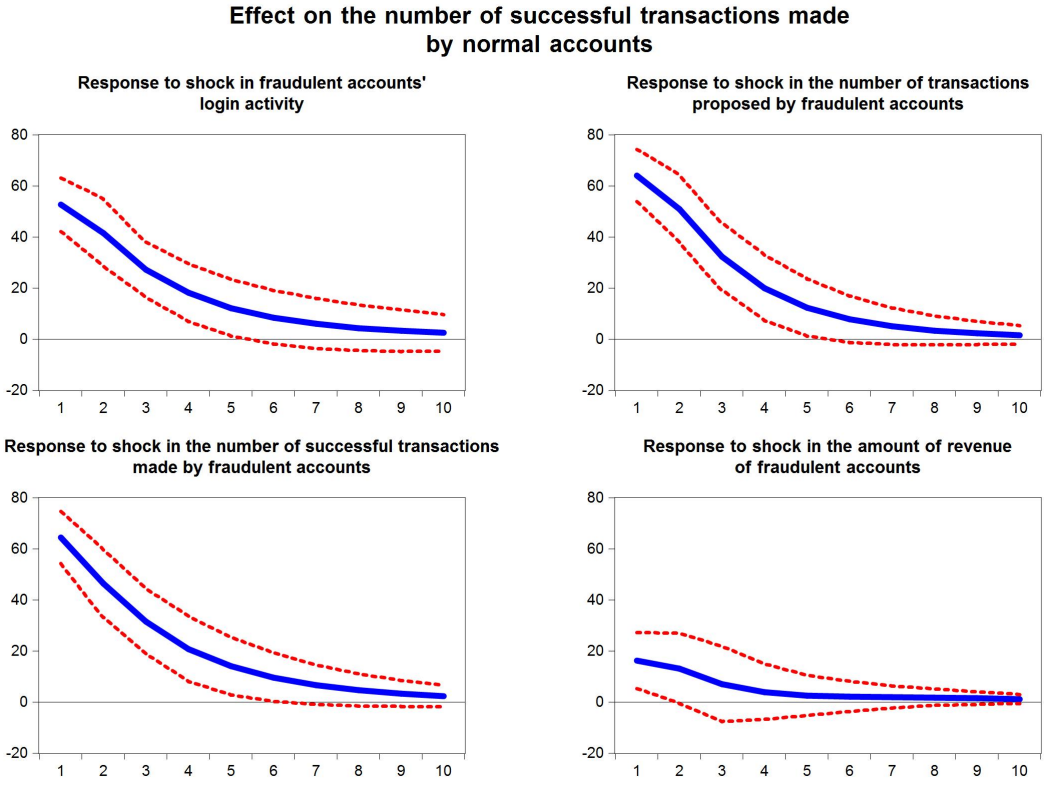


In the figure, the blue curve shows the impact and the red dotted curves represent the one standard deviation confidence interval for the impact.
The x-axis is the number of days. The y-axis is the number of proposed transaction times.
An unexpected positive shock on fraudulent accounts' login times will yield an increase of 208 transaction times on normal accounts' number of proposed transactions. The corresponding elasticity is 0.2209. This positive effect will decrease over time and disappear after 5 days.
An unexpected positive shock on fraudulent accounts' number of proposed transactions will yield an increase of 257 units on normal accounts' number of proposed transactions. The corresponding elasticity is 0.1844. This positive effect will go to zero after 5 days.
An unexpected positive shock on fraudulent accounts' number of successful transactions will lead to 215 more proposed transactions of normal account. The corresponding elasticity is 0.2045. This positive effect will last for 5 days before it becomes insignificant.
An unexpected positive shock on fraudulent accounts' amount of revenue won't yield any significant response on the number of normal accounts' proposed transactions.

same VAR model, the conceptual experiments allow us to separate the direct effects from the overall effects.

In the first conceptual experiment (E1), we allow every variable to have long term effects

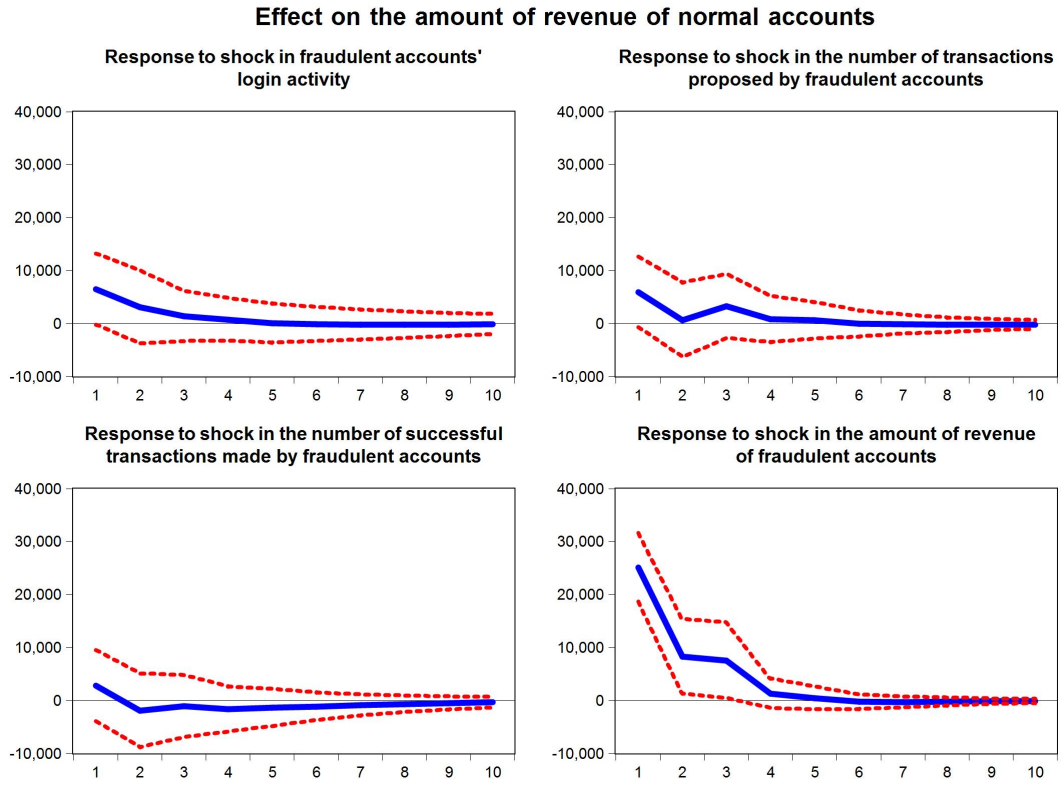
Figure 8: The plots of the generalized impulse response functions for the impact of an unexpected change in fraudulent accounts' variables on the number of successful transaction made by normal accounts.



In the figure, the blue curve shows the impact and the red dotted curves represent the one standard deviation confidence interval for the impact.
The x-axis is the number of days. The y-axis is the number of successful transaction times.
An unexpected positive shock on fraudulent accounts' login times will yield an increase of 53 transaction times on normal accounts' number of successful transactions. The corresponding elasticity is 0.2203. This positive effect will decrease over time and disappear after 5 days.
An unexpected positive shock on fraudulent accounts' number of proposed transactions will yield an increase of 64 units on normal accounts' number of successful transactions. The corresponding elasticity is 0.1797. This positive effect will go to zero after 6 days.
An unexpected positive shock on fraudulent accounts' number of successful transactions will lead to 64 more successful transactions of normal account. The corresponding elasticity is 0.2381. This positive effect will last for 6 days before it becomes insignificant.
An unexpected positive shock on fraudulent accounts' amount of revenue won't yield any significant response on the number of normal accounts' successful transactions.

on its own, and "fraudulent accounts' login activity" to have long term effects on the other variables. This experiment isolates normal accounts' login activity from all the other variables except for itself and the fraudulent accounts' login activity. It therefore allows us to estimate

Figure 9: The plots of the generalized impulse response functions for the impact of an unexpected change in fraudulent accounts' variables on the amount of revenue of normal accounts.



In the figure, the blue curve shows the impact and the red dotted curves represent the one standard deviation confidence interval for the impact.
The x-axis is the number of days. The y-axis is the amount of revenue in 0.01 CHF.
An unexpected positive shock on fraudulent accounts' login times won't yield any significant response on normal accounts' revenue.
An unexpected positive shock on fraudulent accounts' number of proposed transactions won't yield any significant response on normal accounts' revenue.
An unexpected positive shock on fraudulent accounts' number of successful transactions won't yield any significant response on normal accounts' revenue.
An unexpected positive shock on fraudulent accounts' amount of revenue will yield an increase of 25,155 units in the amount of revenue of normal accounts. The corresponding elasticity is 0.2167. This positive effect will become insignificant in about 3 days.

the direct effect of fraudulent accounts' login activity on normal ones' login activity. The second conceptual experiment (E2) adds long term effects of "number of transactions proposed by fraudulent accounts" on the other variables except for "fraudulent accounts' login activity." It allows us to study the direct effect of the number of transactions proposed by fraudulent accounts on that proposed by normal accounts. The third conceptual experiment (E3) adds

Table 8: Elasticity of the retailer in response to the unexpected positive one-standard-deviation shock on fraudulent accounts (Panel A) and normal ones (Panel B).

Panel A: A Shock on Fraudulent Account Levels				
Effects on Retailer	**# Logins**	**# Proposed Transactions**	**# Successful Transactions**	**Revenue**
# Logins	17.79%	—	—	—
# Proposed Transactions	—	27.82%	—	—
# Successful Transactions	—	—	34.9%	—
Revenue	—	—	—	40.25%

Panel B: A Shock on Normal Account Levels				
Effects on Retailer	**# Logins**	**# Proposed Transactions**	**# Successful Transactions**	**Revenue**
# Logins	12.19%	—	—	—
# Proposed Transactions	—	19.77%	—	—
# Successful Transactions	—	—	18.25%	—
Revenue	—	—	—	58.39%

"number of successful transactions made by fraudulent accounts" to have long term effects on the other variables except for "fraudulent accounts' login activity" and "number of transactions proposed by fraudulent accounts". This experiment represents the direct effect of the number of successful transaction made by fraudulent accounts on that made by normal accounts. Finally, the fourth conceptual experiment (E4) adds long term effects of fraudulent accounts' amount of revenue on normal accounts' login activity, number of transactions proposed by normal accounts and fraudulent accounts, and normal accounts' amount of revenue. This conceptual experiment is based on the assumption that purchasing behavior might stimulate one's login activity and encourage customers to propose more trading requests but might not directly lead to more successful transactions. It shows the direct effect of fraudulent accounts' revenue on normal accounts' revenue. Table 9 summarizes the comparisons of direct effects versus indirect effects for these conceptual experiments. The former is estimated by the structural decomposition impulse response functions and the latter equals to the difference between the net and direct effects. The results from the conceptual experiments indicate that in three of the four cases, direct effects have larger effects than indirect effects. The results also show that indirect

effects, which pass through the other variables, also play an important and significant role.

Table 9: Effect separation: the comparison of direct effects (Panel A) versus indirect effects (Panel B).

Panel A: Direct effects **Structural decomposition impulse response functions**				
	A Shock on Fraudulent Account Levels			
Effects on Retailer	**# Logins** **(E1)**	**# Proposed Transactions** **(E2)**	**# Successful Transactions** **(E3)**	**Revenue** **(E4)**
# Logins	114	—	—	—
# Proposed Transactions	—	115	—	—
# Successful Transactions	—	—	48	—
Revenue	—	—	—	24,043

Panel B: Indirect effects **Difference between the generalized impulse response functions and the structural decomposition impulse response functions**				
	A Shock on Fraudulent Account Levels			
Effects on Retailer	**# Logins**	**# Proposed Transactions**	**# Successful Transactions**	**Revenue**
# Logins	45	—	—	—
# Proposed Transactions	—	142	—	—
# Successful Transactions	—	—	16	—
Revenue	—	—	—	1112

Results of the VAR Model: Generalized forecast error variance decomposition

In order to show that fraudulent accounts represent a substantial part of performance variance of normal accounts, we employ the generalized forecast error variance decomposition [45, 52, 41] to show the relative impact of shocks on normal accounts initiated by fraudulent accounts over 10 time periods. Table 10 summarizes the results of the relative impact after 5 and 10 time periods, separately. Based on these results, we can conclude that: (1) fraudulent accounts' activities (number of logins, number of proposed transactions, and number of successful transactions) have substantial impact on normal accounts' activities, and fraudulent accounts' revenue has

significant impact on normal accounts' revenue; (2) The impacts of fraudulent accounts on normal ones are stable over time.

Table 10: Dynamic influence of fraudulent accounts on normal accounts after 5 time periods (Panel A) and 10 time periods (Panel B).

Panel A: Relative impact of shocks initiated by each individual endogenous variable of fraudulent accounts after 5 time periods				
Effects on Normal Accounts (%)	**# Logins**	**# Proposed Transactions**	**# Successful Transactions**	**Revenue**
# Logins	7.542	4.350	4.048	0.667
# Proposed Transactions	5.803	9.186	5.898	0.502
# Successful Transactions	6.877	9.987	9.532	0.603
Revenue	0.745	0.651	0.235	10.374

Panel B: Relative impact of shocks initiated by each individual endogenous variable of fraudulent accounts after 10 time periods				
Effects on Normal Accounts (%)	**# Logins**	**# Proposed Transactions**	**# Successful Transactions**	**Revenue**
# Logins	7.666	4.345	4.102	0.687
# Proposed Transactions	5.846	9.166	5.953	0.499
# Successful Transactions	6.927	9.939	9.575	0.609
Revenue	0.746	0.652	0.273	10.361

Overall, the set of results depicted in the tables and figures indicate that fraudulent accounts consistently have a positive effect on normal accounts. This is true for all three measures of activity, the number of logins, the number of proposed transactions and the number of successful transactions. Furthermore, the effects are persistent, generally remaining significant for up to four to five days after the shock is introduced. The effects of the revenue earned from the fraudulent accounts are less consistent, with the only significant effects being observed on the revenue earned from normal accounts, with the effect lasting about three days. Table 11 summarizes the signs and durations of these effects. These results provide evidence that the activities and revenue of fraudulent accounts on the shopping site has positive effects on the activity levels of the platform as well as on the revenue of the firm (See "Supplementary Information of Chapter 2" for the comparison of VAR model and two other autoregressive models).

Table 11: Signs and durations of the impact of a one standard deviation shock.

	A Shock on Fraudulent Account Levels			
Effects on Normal Accounts Level	**# Logins**	**# Proposed Transactions**	**# Successful Transactions**	**Revenue**
# Logins	+, 5 days	+, 4 days	+, 4 days	Insignificant
# Proposed Transactions	+, 5 days	+, 5 days	+, 5 days	Insignificant
# Successful Transactions	+, 5 days	+, 6 days	+, 6 days	Insignificant
Revenue	Insignificant	Insignificant	Insignificant	+, 3 days

Robustness Checks of the Results

The analysis reported thus far is based on the 5,576 accounts that the retailer identified as fraudulent after completing the manual check and failing to clear the account. All of them were sent the warning email in the first step of the protocol to curb this behavior. However, to test the robustness of our results, we also repeated this analysis using two other, narrower definitions of fraudulent accounts: (1) the subset of 4,345 accounts that were flagged as fraudulent and not only received the email warning but also an in-game warning, and (2) the subset of 3,749 accounts that received all three warnings: email, the in-game warning, plus the confrontation message. The results of the robustness checks are reported in "Supplementary Information of Chapter 2". These analyses show that although the effects of fraudulent accounts on normal ones become less significant, and the durations of the effects become shorter under these narrower definitions of fraudulent accounts, we can still observe the positive impact of fraudulent accounts on normal ones.

Through this study, we sought to validate a key postulate of our theory, that when a longer-term perspective is taken, initial unethical customer actions can sometimes produce positive effects for retailers and peers. Our longitudinal study revealed that at the individual account level, activities of fraudulent accounts, including the number of logins, the number of card trading transactions that they proposed to other customers, and the number of trades that they successfully executed all had positive effects on the activities of normal customers. We found a similar positive pattern of results for the impact of fraudulent accounts' purchasing behavior

on that of normal accounts' purchasing behavior. Our results also indicate that the fraudulent accounts' purchasing behavior had a positive effect on retailer revenue.

2.6 Retailer Response to Beneficial Unethical Customer Behavior

Our theory of retailer response posits that when the retailer realizes the effects of customers' unethical behavior are beneficial for itself and its other customers, in accordance with teleological rather than deontological ethics, it will tend to encourage the behavior. Ethics research also suggests the possibility of differences in response based on ethical ideologies [17]. This line of work distinguishes between ethical relativism and idealism. The chronic trait of ethical relativism is the extent to which the manager rejects universal moral rules because he or she believes that there are many different ways to look at any particular moral issues. Ethical idealism, on the other hand, focuses on moral beliefs that the achievement of desirable consequences will always be obtained when the appropriate actions are performed [17, 3].

Prior research has shown that such people tend to rely on their ethical ideology in how they interpret and respond to questionable actions. Studies show that relativism is less able to predict practical ethical decisions, whereas idealism is more diagnostic (see Davis, Anderson and Curtis 2001, Mudrack and Mason 2013, for a review [12, 40]). Consistent with these findings, we expect that a manager's ethical idealism will predict the decision to ban multiple account customers, but trait relativism will not do so.

Survey-Based Study of Managers' Decision Making

To conduct a test of our predictions, we conducted a survey of retailers. Specifically, we wished to determine the extent to which they would favor retaining the beneficial unethical customers vs. firing them, expecting that a majority would favor retention. We also sought to test our hypothesis that ethical idealism but not relativism would predict the decision to fire beneficial unethical customers. Study Method and Measures. Study participants were 136 US-based

owners or managers of small and medium-sized retail businesses (39.6% female, Mean age = 40 years, SD = 12.0 years) who participated in the online survey in exchange for a $10 e-gift card for Amazon.com. They were recruited using a small business owner e-mail address list to participate in "an academic study of how managers evaluate unethical customer behavior." The mean annual revenue of the businesses was $500,000 – $1 million. Participants were instructed that they would first be given a description of the situation involving a specific type of unethical customer behavior and then asked how they would deal with it. Everyone read the following scenario:

> "Company X is an online firm that sells a variety of products at good prices. Its business model is based on a gamification approach in which customers trade virtual cards with each other. Each offered product has ten different virtual trading cards associated with it. For an offer, the more unique cards customers collect, the greater is the discount they enjoy on its purchase. If they collect all ten cards, they receive the product for free.
>
> Each customer account receives two random cards daily, and can purchase more random cards at a nominal price. They can also trade cards they possess with one another to collect cards for the product that they want to buy.
>
> Because each customer account receives two free cards every day, Company X prohibits its customers from registering multiple accounts on the site. This prohibition is clearly stated in its terms of agreement with customers.
>
> However, a small but significant portion of the customers (around 6%) violate this policy anyway and sign up for multiple accounts on the company's site. By doing this, the customers get more free cards each day, and have more cards to use and trade with others.
>
> Company X's research shows that these customers are also its most engaged and profitable customers. They generate more than a quarter of the firm's revenue and their activity leads to positive effects for the firm and contributes to increased site

activity by other customers who have a single account.

The managers of Company X are unsure about what to do with these customers. On the one hand, they are violating the firm's policy and are engaging in fraudulent behavior. On the other hand, their behaviors benefit both the firm and other customers."

After reading this scenario, participants answered the question "If you were a Senior Manager of Company X, what would you do?" in an open-ended way. Next, they were asked, "If you were a senior manager at Company X, to what extent would you ban these multiple-account customers from the site vs. leave them alone?" Their responses were elicited on a 9-point scale, anchored with 1 = "I would leave these customers alone" and 9 = "I would work very hard to ban these customers from my site." After a short filler task, participants completed Forsyth's (1980) 20-item Ethics Position Questionnaire to assess their idealism and relativism (see "Supplementary Information of Chapter 2"). The two sub-scales showed adequate reliabilities (Idealism α = .86, Relativism α = .83), and were each averaged and mean-centered for use in the analysis. Finally, they completed demographic measures by indicating their age, gender (male = 0, female = 1), education, and annual revenue of their business. We used these four variables as controls in the analysis.

We coded respondents' answers to the open-ended question of what they would do if they were a senior manager into either "Keep the unethical customers" or "Get rid of the unethical customers." Responses in the "keep customers" category included:

"Allow them to continue. As management has created the rules they have the latitude by which to enforce or make exceptions. As no harm is being done, except potentially to the company who is exposed to more giveaways, this is within their purview."

"I would not do anything to discourage or obstruct this behavior. If anything, I might either change the company's Terms of Agreement, or even try to implement changes that formalize what these 6% of users are already doing. For example,

increasing the number of cards that customers receive based on the age or activity level of their account."

"I wouldn't do anything but I would keep an eye on the percentage of these customers."

Responses in the "get rid of customers" category included:

"I would stop those unethical actions that violates the terms of the agreement."

"I would give them warning and terminate them if they keep this activity."

"I would confront the so called violators to avoid lawsuits within the company.

"I would either modify the policy for everyone, or enforce the current policy and penalize the people not following the rules."

Of the sample, 19.9% of respondents indicated that they would get rid of the unethical customers, whereas the majority or 80.1% of respondents indicated that they would keep the unethical customers. The two dimensions of ethical ideologies, idealism and relativism, were uncorrelated to each other (r = .012, p = .89). We ran hierarchical multiple regressions in which the respondents' stated extent to which multiple-account customers would be banned vs. left alone was regressed on the controls, relativism, and idealism. The controls were entered in the first block, relativism was entered in the second block, and idealism in the third block. Results showed that the controls explained 15.7% of the variance in the decision to ban vs. leave multiple account customers alone with age (β = .187, p = .027), gender (β = .264, p = .002), and education (β = -.176, p = .035) all predicting this variable significantly. Adding relativism did not increase explained variance at all (β = -.003, p = .973) in the second block, but idealism was a significant predictor (β = .258, p = .004) with explained variance of 20.9% (ΔR^2 = 33% over the baseline). As expected, greater idealism of the business owner was associated with greater insistence on banning multiple-account customers from the site, but relativism did not affect the decision.

These findings show that consistent with our theorizing, a majority of managers support keeping the customers even though their behavior is unethical. However, the ethical ideology

of the managers involved plays a role such that ethical idealists are more likely to insist on the firm deontological stance of banning multiple-account customers.

2.7 General Discussion

Research Contributions

The extant retailing literature and practice adopt a fairly rigid and narrow perspective on unethical customers that is focused on loss prevention and deterrence. Against this backdrop, we introduced a theoretical framework that shifts the focus away from the rightness or wrongness of the customer's action, and towards understanding its' consequences, especially over the longer-term, before deciding how to respond to it. While acknowledging that many unethical customer behaviors are too serious to warrant leniency, we argue that in today's digital, information-dense, privacy-conscious, and customer relationship-oriented environment, many retailers have adopted conservative, tradition-based policies of what is allowed and not allowed by customers [39]. Even when a customer violates such policies, thus committing an ethical transgression, its consequences may not always negative. In fact, when longer-term implications of such actions are considered, retailers and non-transgressing customers may both benefit as was the case with the retailer in our study.

We believe our broader theoretical perspective on retailer response introduces pragmatism, based on a teleological ethical perspective, into the retailer's decision making process. It is likely to benefit retailers of all stripes in re-thinking how they conceive of, and deal with, their customers' unethical behavior. The theory introduced here also opens the door for a deeper study of retailers' policies directed towards customers and their effects on customer behavior and retailer outcomes. We can easily imagine that similarly beneficial effects could occur to retailers from return policy violators [56], at least from some customer segments; but our speculation needs to be validated through future research. Another issue worth exploring is to specify the conditions that lead unethical behaviors to produce positive retailer outcomes such as those we postulated and found.

Why did unethical customer behavior have positive effects?

Our study's setting has at least three characteristics that are likely to have contributed to the positive effects of unethical customer behavior. It is worth discussing these conditions because they provide a better understanding of where else such effects may occur, and may help to deepen our proposed theory in future research.

The first condition is that the customer's unethical action concerns an information good that the retailer produced and distributed with negligible variable costs [1]. Specifically, the retailer incurred minimal incremental costs for registering multiple accounts of unethical customers, and for generating additional trading cards for them because of their multiple logins[2] . Another example where this condition may obtain is a customer rewards program that offers points to enhance a customer's status within a social media platform. Here too, it is likely that multiple accounts opened by a customer would enhance activity and produce more revenue-generating opportunities (e.g., through advertising) without adding commensurate costs to the marketer.

The second reason for beneficial effects is the possibility that the customer's initial unethical action of registering multiple accounts propelled the customer's relationship with the retailer into a different, more active trajectory. With greater incentives (multiple daily free cards) and resulting benefits accrued from trading cards and making discounted purchases, having multiple accounts led to greater activity, which accumulated over time. Thus, the initial unethical action directly encouraged latter customer engagement, and generated greater revenues for the retailer. It is noteworthy that the unethical behavior concerned only initial registration; later trading and purchase behavior was in line with that of other ethical customers (albeit with multiple accounts). The active trading of the fraudulent customer boosted the amount of login and trading activity of other customers, benefitting their relationships with the retailer as well. For online retailers reliant on advertising-based revenue models, greater customer engagement will have additional payoffs.

[2]We note that the rarity of the generated cards was determined independently and was not influenced in any way by fraudulent accounts. The presence of fraudulent accounts simply impacted which customers won the rare cards, but did not result in more rare cards being generated.

The third reason for positive effects is likely to be the fact that there was no obvious harm to any identifiable victim — either other customers or the retailer itself — from the unethical customer's actions [19]. As mentioned earlier, the most significant potential harm from multiple registrations is that it reduced ethical customers' chances of winning rare cards. But the results indicate that: (a) either the customers did not notice this, (b) found it to be unimportant, or (c) the greater number of cards available for trade on the site made up for any perceived detriment, or a combination of the above[3].

2.8 Conclusion

A more practical and actionable compromise that bridges the gap between these two philosophical perspectives is that the retailer should change its user policies so that an individual customer can maintain and utilize more than one user profile on the site without the need for multiple telephone numbers or email addresses. This is consistent with the approach used by sellers of information goods such as Netflix which allows its customers to readily share their account with others and create up to five different user profiles and customize preferences within each profile [9]. For this retailer, such flexibility would lead to the customer's behavior that is currently seen as surreptitious and ethically questionable to be clearly seen as morally acceptable under both deontological and teleological philosophies.

To conclude, our research compellingly uncovers the counter-intuitive phenomenon of unethical consumer behavior having predominantly positive consequences for the retailer and for other customers. It empirically illustrates the observation made by Donaldson and Dunfee (1994, p. 258) [15], that "the ethical norms must be contoured to the rules of the specific economic practices and the notions of fairness of participants." Consequently, there is a strong need to consider unethical consumer behaviors in retailing contexts in more nuanced and balanced ways and to devise solutions that are equally nuanced and lead to the best possible outcomes for customers and for retailers as a whole, even when such actions do not fall strictly within the

[3]We note that the retailer did not publicize these customers' fraudulent behavior at any time during the study period (or since then) to the best of our knowledge. All communications and warnings to unethical customers were conducted on an individual basis and discreetly.

parameters set forth by the influential moral philosophies.

References

[1] Bakos, Y. & Brynjolfsson, E. Bundling information goods: Pricing, profits, and efficiency. *Management science* **45**, 1613–1630 (1999).

[2] Bamfield, J. A. Shopping and crime. In *Shopping and Crime*, 1–10 (Springer, 2012).

[3] Barnett, T., Bass, K. & Brown, G. Ethical ideology and ethical judgment regarding ethical issues in business. *journal of Business Ethics* **13**, 469–480 (1994).

[4] Bellur, V. V. Shoplifting: Can it be prevented? *Journal of the Academy of Marketing Science* **9**, 78–87 (1981).

[5] Bentham, J. *An introduction to the principles of morals and legislation* (Oxford: Clarendon Press, 1789).

[6] Blanco, C. *et al.* Prevalence and correlates of shoplifting in the united states: results from the national epidemiologic survey on alcohol and related conditions (nesarc). *American Journal of Psychiatry* **165**, 905–913 (2008).

[7] Bohns, V. K., Roghanizad, M. M. & Xu, A. Z. Underestimating our influence over others' unethical behavior and decisions. *Personality and Social Psychology Bulletin* **40**, 348–362 (2014).

[8] Bolton, R. J. & Hand, D. J. Statistical fraud detection: A review. *Statistical science* 235–249 (2002).

[9] Chaey, C. Now you can have multiple user profiles on one netflix account (2013). URL `http://www.fastcompany.com/3015138/fast-feed/now-you-can-have-multiple-user-profiles-on-one-netflix-account`.

[10] Conan, N. Busted: What happens when shoplifters get caught? (2012). URL `http://www.npr.org/2012/11/15/165218744/busted-what-happens-when-shoplifters-get-caught`.

[11] Cox, D., Cox, A. D. & Moschis, G. P. When consumer behavior goes bad: An investigation of adolescent shoplifting. *Journal of consumer research* **17**, 149–159 (1990).

[12] Davis, M. A., Andersen, M. G. & Curtis, M. B. Measuring ethical ideology in business ethics: A critical analysis of the ethics position questionnaire. *Journal of Business Ethics* **32**, 35–53 (2001).

[13] Dekimpe, M. G. & Hanssens, D. M. Sustained spending and persistent response: A new look at long-term marketing profitability. *Journal of Marketing Research* 397–412 (1999).

[14] Dickey, D. A. & Fuller, W. A. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association* **74**, 427–431 (1979).

[15] Donaldson, T. & Dunfee, T. W. Toward a unified conception of business ethics: Integrative social contracts theory. *Academy of management review* **19**, 252–284 (1994).

[16] Erat, S. & Gneezy, U. White lies. *Management Science* **58**, 723–733 (2012).

[17] Forsyth, D. R. A taxonomy of ethical ideologies. *Journal of Personality and Social psychology* **39**, 175 (1980).

[18] Gino, F. Understanding ordinary unethical behavior: why people who value morality act immorally. *Current opinion in behavioral sciences* **3**, 107–111 (2015).

[19] Gino, F., Shu, L. L. & Bazerman, M. H. Nameless+ harmless= blameless: When seemingly irrelevant factors influence judgment of (un) ethical behavior. *Organizational Behavior and Human Decision Processes* **111**, 93–101 (2010).

[20] Gino, F., Ayal, S. & Ariely, D. Self-serving altruism? the lure of unethical actions that benefit others. *Journal of economic behavior & organization* **93**, 285–292 (2013).

[21] Granger, C. W. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society* 424–438 (1969).

[22] Hanssens, D. M. Market response, competitive behavior, and time series analysis. *Journal of Marketing Research* 470–485 (1980).

[23] Hanssens, D. M., Parsons, L. J. & Schultz, R. L. *Market response models: Econometric and time series analysis*, vol. 12 (Kluwer Academic Publishers, 2001).

[24] Harris, L. C. Fraudulent return proclivity: an empirical analysis. *Journal of Retailing* **84**, 461–476 (2008).

[25] Harris, L. C. Fraudulent consumer returns: exploiting retailers' return policies. *European Journal of Marketing* **44**, 730–747 (2010).

[26] Hoekman, A. R. *Preventing Shrink: Life in Loss Prevention* (New York: CreateSpace, 2015).

[27] Hunt, S. D. & Vitell, S. J. The general theory of marketing ethics: A retrospective and revision. *Ethics in Marketing (Irwin Inc., Homewood, IL)* 775–784 (1993).

[28] Jackson, J. Telling the truth. *Journal of medical ethics* **17**, 5–9 (1991).

[29] Jones, T. M. Ethical decision making by individuals in organizations: An issue-contingent model. *Academy of management review* **16**, 366–395 (1991).

[30] Jones, T. M., Felps, W. & Bigley, G. A. Ethical theory and stakeholder-related decisions: The role of stakeholder culture. *Academy of Management Review* **32**, 137–155 (2007).

[31] Kant, I. *Fundamental principles of the metaphysics of morals* (London: Pearson, 1785).

[32] Kim, J., Kim, J.-E. & Park, J. Effects of cognitive resource availability on consumer decisions involving counterfeit products: The role of perceived justification. *Marketing Letters* **23**, 869–881 (2012).

[33] Levine, E. E. & Schweitzer, M. E. Are liars ethical? on the tension between benevolence and honesty. *Journal of Experimental Social Psychology* **53**, 107–117 (2014).

[34] Luo, X., Raithel, S. & Wiles, M. A. The impact of brand rating dispersion on firm value. *Journal of Marketing Research* **50**, 399–415 (2013).

[35] Lütkepohl, H. *New introduction to multiple time series analysis* (Berlin: Springer Science & Business Media, 2007).

[36] Mark, T., Lemon, K. N., Vandenbosch, M., Bulla, J. & Maruotti, A. Capturing the evolution of customer–firm relationships: How customers become more (or less) valuable over time. *Journal of Retailing* **89**, 231–245 (2013).

[37] Mazar, N., Amir, O. & Ariely, D. The dishonesty of honest people: A theory of self-concept maintenance. *Journal of marketing research* **45**, 633–644 (2008).

[38] Mazar, N. & Ariely, D. Dishonesty in everyday life and its policy implications. *Journal of public policy & Marketing* **25**, 117–126 (2006).

[39] Miyazaki, A. D. & Fernandez, A. Internet privacy and security: An examination of online retailer disclosures. *Journal of Public Policy & Marketing* **19**, 54–61 (2000).

[40] Mudrack, P. E. & Mason, E. S. Ethical judgments: What do we know, where do we go? *Journal of Business Ethics* **115**, 575–597 (2013).

[41] Nijs, V. R., Srinivasan, S. & Pauwels, K. Retail-price drivers and retailer profits. *Marketing Science* **26**, 473–487 (2007).

[42] Nyberg, D. The morality of everyday activities: Not the right, but the good thing to do. *Journal of Business Ethics* **81**, 587–598 (2008).

[43] Pauwels, K. How dynamic consumer response, competitor response, company support, and company inertia shape long-term marketing effectiveness. *Marketing Science* **23**, 596–610 (2004).

[44] Pauwels, K. How retailer and competitor decisions drive the long-term effectiveness of manufacturer promotions for fast moving consumer goods. *Journal of Retailing* **83**, 297–308 (2007).

[45] Pesaran, H. H. & Shin, Y. Generalized impulse response analysis in linear multivariate models. *Economics letters* **58**, 17–29 (1998).

[46] Phillips, P. C. & Perron, P. Testing for a unit root in time series regression. *Biometrika* **75**, 335–346 (1988).

[47] Reynolds, K. L. & Harris, L. C. Dysfunctional customer behavior severity: an empirical examination. *Journal of retailing* **85**, 321–335 (2009).

[48] Ruedy, N. E., Moore, C., Gino, F. & Schweitzer, M. E. The cheater's high: The unexpected affective benefits of unethical behavior. *Journal of Personality and Social Psychology* **105**, 531 (2013).

[49] Sennewald, C. A. & Christman, J. H. *Retail crime, security, and loss prevention: an encyclopedic reference* (Oxford, UK: Butterworth-Heinemann, 2008).

[50] Shu, L. L., Gino, F. & Bazerman, M. H. Dishonest deed, clear conscience: When cheating leads to moral disengagement and motivated forgetting. *Personality and Social Psychology Bulletin* **37**, 330–349 (2011).

[51] Sims, C. A. Macroeconomics and reality. *Econometrica: Journal of the Econometric Society* 1–48 (1980).

[52] Srinivasan, S., Pauwels, K. & Nijs, V. Demand-based pricing versus past-price dependence: A cost-benefit analysis. *Journal of Marketing* **72**, 15–27 (2008).

[53] Trusov, M., Bucklin, R. E. & Pauwels, K. Effects of word-of-mouth versus traditional marketing: findings from an internet social networking site. *Journal of marketing* **73**, 90–102 (2009).

[54] Vitell, S. J. Consumer ethics research: Review, synthesis and suggestions for the future. *Journal of business ethics* **43**, 33–47 (2003).

[55] Wilkes, R. E. Fraudulent behavior by consumers. *Journal of Marketing* **42**, 67–75 (1978).

[56] Wood, S. L. Remote purchase environments: The influence of return policy leniency on two-stage decision processes. *Journal of Marketing Research* **38**, 157–169 (2001).

[57] Zichermann, G. & Linder, J. *Game-based marketing: inspire customer loyalty through rewards, challenges, and contests* (Hoboken, New Jersey: John Wiley & Sons, 2010).

3 Fraudulent Behavior and Statistical Fraud Detection Techniques: A Review[4]

Abstract

Fraudulent behavior is a serious problem all over the world and occurs in practically every business. It has become a major phenomenon representing substantial amounts of losses. From banks to E-commerce, fraudsters keep on finding failures in the existing systems, which makes fraud detection techniques an important and highly necessary field of research. In this paper, the authors review the literature in standard economic, behavioral ethics, psychology, and neuroscience on unethical behavior in order to understand its causes and consequences. Subsequently, we focus on studies related to ordinary unethical behavior. Furthermore, various fraud detection techniques are reviewed, in which we especially highlight the application of network science to fraud detection. Managerial implications as well as the possible future directions of the above-mentioned researches are discussed in the end.

3.1 Introduction

> "Your Honor, for many years up until my arrest on December 11, 2008, I operated a Ponzi scheme through the investment advisory side of my business, Bernard L. Madoff Securities LLC, which was located here in Manhattan, New York at 885 Third Avenue. I am actually grateful for this first opportunity to publicly speak about my crimes, for which I am so deeply sorry and ashamed. As I engaged in my fraud, I knew what I was doing was wrong, indeed criminal. When I began the Ponzi scheme I believed it would end shortly and I would be able to extricate myself and my clients from the scheme. However, this proved difficult, and ultimately

[4] **Author Statement:** This is a working paper together with Manuel Schnurrenberger, Alexandre Guinand, and René Algesheimer. It has not yet been submitted to any journal. Zhao Yang is the 1^{st} and corresponding author, Manuel Schnurrenberger is the 2^{nd} author, Alexandre Guinand is the 3^{rd} author, and René Algesheimer is the 4^{th} author.

> impossible, and as the years went by I realized that my arrest and this day would inevitably come. I am painful aware that I have deeply hurt many, many people, including the members of my family, my closest friends, business associates and the thousands of clients who gave me their money. I cannot adequately express how sorry I am for what I have done. I am here today to accept responsibility for my crimes by pleading guilty and, with this plea allocution, explain the means by which I carried out and concealed my fraud." (Bernard L. Madoff)

When it comes to fraud, people are likely to think about Enron, WorldCom, and Bernard Madoff, etc. This is simply due to the fact that media reports only the most sensational cases and huge corporate scandals. In the above-mentioned cases, Enron and WorldCom are two well-known corporate scandals [100, 115], and Bernard Madoff together with his Ponzi Scheme is a famous sensational case [114]. All of them have had damaging consequences to the society.

Indeed, fraudulent behavior is a serious problem all over the world and occurs in practically every business. According to the global fraud report of 2015/2016 [85], despite that companies make greater and more sophisticated efforts to combat fraud, it remains a big business threat that cannot be completely eliminated. This annual report claimed that fraud has continued to increase, with three quarters of companies reporting they have fallen victim to a fraud incident within the past year, an increase of 14 percentage points from just three years ago; and the number of businesses suffering a financial loss as a result of fraud has also increased, from 64% in the previous survey period to 69% this year. It also reveals three key trends that firms feel more vulnerable to fraud; the globalization of business increases fraud risk; and the biggest fraud threat to companies comes from within.

In many extreme cases, the responsible persons and what they have done are unforgivable. Taken Bernard Madoff as an example, in his plea allocution, he claimed that he knew what he was doing was criminal and he felt deeply sorry and ashamed. He was sentenced to 150 years in prison for turning his wealth management business into a massive Ponzi scheme that lasted about 20 years and involved around $65 Billion. However, apart from these extreme cases, "ordinary" fraud behaviors, such as overstatement of performance at work, return abuse,

and office supply scams, are actually more frequency and pervasive. These kinds of fraud behaviors are more acceptable and are hardly reported by any public media. Contradictory to our common feeling that bad behavior is tied to bad character, or so called bad people do bad things, unethical actions can also be committed by people who value and care about morality but behave unethically when faced with an opportunity to cheat [55]. For instance, United States retailers are losing $60 Billion a year to shrinkage in 2015, up from $57 Billion in 2014, and employee theft is the single biggest cause of loss to retailers [86]. Another example is about peer-to-peer music downloading. A survey conducted by Fred von Lohmann in 2004 claimed that 88% of children between 8 and 18 years understood that peer-to-peer music downloading is illegal and unethical, but 56% of the same children admitted to continuing doing so [90].

Although many companies are victims of fraudulent behavior, their attitudes toward fraud are also twisted. A recent survey with more than 2800 senior executives from 62 countries has found that a significant minority of executives continue to justify unethical behavior to improve a company's performance. When presented with a series of options, more than one-third would be willing to justify inappropriate conduct in an economic downturn, while almost half would justify such conduct to meet financial targets or safeguard a company's economic survival [40]. The high percentage of managers that would help the company to hide or justify unethical behavior to reach certain monetary targets is alarming for the societal welfare. Another survey conducted in our first project with 138 owners of small and medium-sized businesses in the US has obtained similar conclusions. Given the scenario that unethical customers might have predominantly positive consequences for the retailer and for other customers, more than 80% of respondents are fine with keeping the unethical customers [See Chapter 2 for detail].

Overall, unlike sensational fraudulent cases and huge corporate scandals, ordinary unethical behavior, justified by normal people like you and me, or executives in various companies, is more accepted and common than one thinks. Although received little attention by public media, it has a tremendous impact on economics, marketing activities, as well as our daily life.

The rest of the paper is organized as follows. The definition of fraud and its related terms will be given in section 3.2. After that, in section 3.3, we try to capture the most important the-

ories and concepts in unethical behavior by reviewing literatures in standard economic, behavioral ethics, psychology, and neuroscience in order to understand the causes and consequences of fraudulent behavior. In section 3.4, we especially focus on studies in ordinary unethical behavior committed by ordinary people who care about morality. Furthermore, various statistical fraud detection techniques are reviewed in section 3.5. We will then highlight the application of network science in fraud detection in section 3.6. Possible managerial implications are discussed in section 3.7. Finally, in section 3.8, there is a brief wrap-up together with an outlook on future steps at the end of the paper.

3.2 Definitions of fraud, cheating, dishonest, immoral, and unethical behavior

Fraud, wrongful or criminal deception intended to result in financial or personal gain, comes from Old French "fraude" and Latin "fraus" [101]. Fraud itself can be a civil wrong, a criminal wrong, or it may cause no loss of money, property or legal right but still be an element of another civil or criminal wrong [140]. Fraud can also be defined from psychological point of view, for instance, Olsen has claimed: "Fraud is a human endeavor, involving deception, purposeful intent, intensity of desire, risk of apprehension, violation of trust, rationalization, etc." [110]. From a legal point of view, fraudulent behavior is prohibited in many countries. For instance, the Article 146 of the Swiss Criminal Code indicates that "any person who with a view to securing an unlawful gain for himself or another wilfully induces an erroneous belief in another person by false pretences or concealment of the truth, or wilfully reinforces an erroneous belief, and thus causes that person to act to the prejudice of his or another's financial interests, is liable to a custodial sentence not exceeding five years or to a monetary penalty. If the offender acts for commercial gain, he is liable to a custodial sentence not exceeding ten years or to a monetary penalty of not less than 90 daily penalty units" [130].

A synonym for fraud is cheating [102]. Cheating is committing fraud and/or deception on a record, report, paper, computer assignment, examination, or any other course or field

placement assignment [136]. Both fraud and cheating are different forms of unethical behavior [42]. Unethical behavior is a much broader category of wrongdoing than fraudulent behavior and cheating. For instance, Jones has described unethical behavior as behavior that is "either illegal or morally unacceptable to the larger community" [73]. It is an action that falls outside of what is considered morally right or proper for a person, a profession or an industry [143]. Therefore, unethical behavior is not necessary to be intended, which makes it different from fraudulent behavior. As always mentioned together, dishonest and immoral are synonyms for unethical [133].

3.3 Development of the research of unethical behavior

Without loss of generality, in this section, we will provide an overview on why people commit unethical behavior. In literature there are two main research streams that try to explain the mechanisms of unethical behavior on an individual level [89]. The first research stream is referred to as the standard economic perspective. It has its roots in the late Middle Age and reached its zenith in the 20th century, especially in the field of crime and punishment, to explain when someone commits an unethical act. In contrast to the purely economic view, there exists ample evidence from other academic disciplines such as psychology, behavior economics and neuroscience, that in addition to external reward mechanisms, there are internal reward mechanisms that influence the individual decision-making process and behavior as well. Hence, parallel to the standard economic perspective, a second major research stream, called the psychological perspective, evolved in the second half of the 20th century. After these two main research streams are reviewed, some insights from emerging approaches such as behavior economics and neuroscience are examined.

3.3.1 Standard economic perspective

The standard economic perspective is rooted in the theoretical and philosophical concepts of Thomas Hobbes and Adam Smith [71, 127]. This perspective assumes that a human being acts

fully rational, selfish and maximizes its own payoffs. This rational human being – the so-called homo economicus – knows always what he wants and is able to choose the best possible option available. Becker transferred the idea of homo economicus to the area of crime and punishment [18, 19]. In his model of the "economics of crime", the goal was to minimize criminal acts because at the time he wrote that theory, crime had grown rapidly and reduced social welfare. He stated that crime could be decreased when the probability of being caught and the form and size of the punishment are adjusted by the policy of the state. In addition, Becker already noticed that crime could be effectively decreased with the certainty of being caught but not with the size and form of the penalty. This finding is supported by experiments of Nagin and Pogarsky in 2003 [98]. Allingham and Sandmo developed the model of Becker further and used it to explain why an individual decides to evade taxes: If an evader gets caught at a certain time period, all his preceding tax evasions are discovered as well. This, in turn, makes the evader more risk averse [7]. This view on why people engage in tax evasion can also be applied in the field of insurance fraud and in general for computing whether someone commits fraudulent behavior or not [36, 43, 89]. For instance, an individual would consider only three external factors when they would pass a gas station shop. First, the expected amount of money the individual would gain from robbing the shop. Second, the probability of been caught while robbing the station shop. And third, the magnitude of punishment if caught [89]. The decision to commit the unethical act only depends on the expected external benefits (e.g. money gained, or getting a better position at work) and the expected external costs (e.g. paying a fine, or losing a job). Assuming the three above-mentioned external factors are known in the gas station shop example, an individual chooses to rob the shop in case he gets a positive utility out of the act. The decision-making process is a deliberate, conscious act made by the individual. However, because the decision depends only on the external cost-benefit analysis, the individual's internal thoughts and processes itself have no influence on the outcome of his decision [90].

In the course of time, the standard economic perspective was incrementally extended. In the field of crime and punishment theory, the economical perspective was extended with psychological factors, because the assumptions of homo economicus were not sufficient to explain human

behavior [70, 80]. Critics claim that given an opportunity to commit crime, individuals with low self-control will do so [70, 128]. In the field of tax evasion, experiments give evidence that subjects respond not only to probabilities as Allingham and Sandmo assumed [7] but also to the context provided such as the perception of the fairness of tax system or trust in government [126, 81].

An overview of articles mentioned in this section can be found in Table 12. The articles are sorted based on appearance in the manuscript.

Table 12: Overview of the articles mentioned in Section 3.3.1: Standard economic perspective.

Paper	**Key points**
Hobbes (1968), Smith and Nicholson (1887)	Homo economicus
Becker (1962, 1968)	Introduce the concept of "Homo economicus" to the area of crime and punishment
Nagin and Pogarsky (2003)	Prove that crime could be decreased with the certainty of being caught but not with the size and form of the penalty
Allingham and Sandmo (1972)	Extend Becker's model to explain tax evasion
Derrig (2002), Farashah and Estelami (2014)	Apply Allingham and Sandmo's explanation in the filed of insurance fraud
Mazar and Ariely (2006)	Describe a general way of computing whether someone commits fraudulent behavior or not based on cost-benefit analysis
Hirschi and Gottfredson (2001), Kleemans (2013), Smith (2004), Slemrod (2007), Kleven et al. (2011)	Introduce psychological factors to extend "homo economicus". For instance, low crime; self-control in perception of fairness in tax evasion

3.3.2 Psychological perspective

As has just been mentioned, from the psychological perspective, additionally to the external cost-benefit consideration, other important psychological factors enter the consideration of an individual when deciding to act (un)ethical [89].

Previous theories of ethical decision making in organizations have tended to emphasize

either the individual or situational factors in explaining (un)ethical behavior [35, 138]. The interactionist model from Trevino, however, tries to integrate both the individual and the situational factors that lead to certain behaviors [134]. This model is based on individual moderators such as ego strength, field dependence and locus of control, and on situational factors such as the imitate job context, organizational culture and the characteristics of the work.

Another important framework for the analysis of decision-making comes from Ferrel and Gresham [51]. They present an integrated contingency framework explaining (un)ethical decision-making with the help of 3 factors: individual factors (knowledge, attitudes), significant others (role set configuration) and opportunity (reward, punishment by policy makers, code of conduct). An extended framework has later been presented by Jones who added that the moral issue itself can also affect the decision-making process [73]. All three factors in this integrated contingency framework influence the decision of an individual to behave ethically or not. For instance, Shariff and Norenzayan showed that people who believe in a more fearsome punishing God, than a compassionate one, cheated less in anonymous situations than others [122]. Asch conducted experiments about conformity behavior in groups, while Milgram investigated the conditions that produce obedience to authority [10, 95]. With their experiment "the good Samaritan", Darley and Batson concluded that situational variables are significant predictors of (un)ethical behavior [35]. The findings of Doris also support the hypothesis that situational factors are even more important in predicting behavior than individual dispositions [38].

Except for the above-mentioned two models, the moral development model is also an important theory in the decision-making process. The idea behind this theory traces back to Piaget [107]. Kohlberg and Hersh later refined the theory further [82]. This theory is a model with 6 stages and 3 levels. The 6 stages are: (1) The punishment-and-obedience orientation, (2) The instrumental-relativist orientation, (3) The interpersonal concordance orientation, (4) The "law and order" orientation, (5) The social-contract, legalistic orientation, and (6) The universal-ethical-principle orientation. And the three levels are: (1) Pre-conventional level, (2) Conventional level, and (3) Post-conventional, autonomous level. The moral development model

assumes that people can progress to a higher and more adequate stage of moral reasoning, if they have the psychological capacity. The interaction with one's environment is the key factor of the development of moral reasoning. Not everyone will reach the highest stage namely stage 6 and hence, not everyone is able to judge ethically in a complex context and therefore, not everyone acts the same way in ethical issues. The moral development model is extended by Rest et al. [113]. It consists of 3 schemas: (1) the personal interest schema, (2) the maintaining schema, and (3) the post-conventional schema.

In contrary to the moral development theory, the social cognitive theory states that behavior is often determined by automatic processes. Therefore deliberative moral reasoning is not always needed. This theory is originally from Bandura [11]. It views people as interactive agents, who are proactive, self-reflecting and self-regulating rather than just reactive organisms shaped by external forces (environment).

An important concept in the social cognitive theory is self-regulatory, or self-control. According to Gino et al., self-control is the internal process that helps individuals to resist short-term temptations and to achieve long-term goals [61]. It is also called "moral muscle" and it can be depleted when individuals continue to exert self-control without rest [92]. A good example to explain self-control, in the context of unethical behavior, is overstating performance at the work place. Overstating performance offers individuals short-term benefits (e.g. monetary gains) but at the same time it can cause long-term costs (e.g. bad reputation or lower social acceptance). In such a dilemma, individuals have to weigh two opposing forces: the desire of self-interested maximization and the desire to maintain a positive moral self-image [61]. Experiments provided evidence that self-control depletion promotes unethical behavior and impairs individual's ability to recognize that their behavior is unethical [92, 61]. A lack of sleep could diminish self-control, and in turn, reduces the inhibition of unethical behavior [14].

According to the social cognitive theory, moral reasoning – or in other words: what one ought to do – is translated into actions by self-regulatory mechanisms. In the self-regulatory process, individuals monitor their actions and judges based on their moral standards. If the actions do not meet their internal moral standards, individuals regulate the actions through self-

sanctions [12]. Internal standards emerge from a concept called moral identity. Moral identity is the cognitive pattern a person holds about his or her moral character [9]. It is a powerful source of moral motivation due to the fact that people would like to maintain self-consistency, or a positive self-view [89, 9]. Furthermore, moral disengagement is a set of cognitive mechanisms that cause a deactivation of moral self-regulation and allows people to make unethical decisions more easily [37]. Based on the social cognitive theory, in the early stage of development of an individual, behavior is largely regulated by external factors and social sanctions. Later, during the course of socialization, individuals adopt moral standards that serve as an ethical guide and are a major source for self-sanctions [12, 13].

Money has been found to play an important role in decision-making also [59, 60, 137]. For instance, it has been shown in some experiments that the mere presence of substantial wealth leads to perceptions of inequity among employees in an organization [60]. This perception of negative inequity induces a feeling of envy and motivates individuals to commit fraud such as a deceptive overstatement of performance at the workplace.

People sometimes behave unethically to benefit their group rather than for self-interests. Recent results from experiments of Gino, Ayal, and Ariely show, if individual's unethical behavior could benefit others, the level of individual cheating increased [56, 57]. Furthermore, the level of cheating increased even if cheating only benefited the others and not the self. These findings contradict the economic perspective, where only self-interested external rewards are considered. Another study of Thau et al. reveals that employees who believed they were at risk of social exclusion of the working group engaged more in pro-group unethical behavior than others [132].

Except for the studies of ethical decision-making, there are other frameworks in understanding the causes of unethical behavior. For instance, the fraud triangle framework explains the motivation of an individual to commit fraudulent behavior with three elements: Perceived unshareable financial need, perceived opportunity, and rationalization of fraudulent behavior. This framework originated from Donald Cressey's hypothesis: "Trusted persons become trust violators when they conceive of themselves as having a financial problem which is non-shareable,

are aware this problem can be secretly resolved by violation of the position of financial trust, and are able to apply to their own conduct in that situation verbalizations which enable them to adjust their conceptions of themselves as trusted persons with their conceptions of themselves as users of the entrusted funds or property" [34]. According to Ramamoorti, incentives and perceived pressure motivate to commit fraud at the first place [111]. For example, an individual has some financial problems that he is not able to solve through legal means, so he starts to consider to commit fraud to overcome the pressure [3]. The opportunity includes the assessment of being caught. The last element, rationalization of fraudulent behavior is the new element this framework evokes. Since then, a large body of research has studied the concept of rationalization [15, 131, 54, 65, 121]. The need to rationalize unethical behavior is rooted in the cognitive dissonance. A cognitive dissonance emerges from the gap between the own perceptions of being ethical and the related unethical actions or behaviors one takes [52]. According to the "fraud triangle" an individual most likely commits fraud, if he has a good motive/incentive and/or is under pressure; if the opportunity of being caught is negligible; and if the fraudster is able to rationalize his actions to himself [78].

Although the "fraud triangle" is a useful tool to examine why people commit fraud, critics argue that pressure and rationalization cannot be observed empirically and important factors like the capability of the fraudster is ignored [78]. To iron those faults out, many different researchers expanded the "fraud triangle" from Cressey. Albrecht, Howe, and Romney introduced the "fraud scale model" [6]. This model replaces rationalization through personal integrity because integrity can be observed and measured. Wolfe and Hermanson developed the "fraud diamond model", where they added the factor the fraudster's capabilities [141]. The authors believed that a lot of fraud would not have been committed, if the person did not have the right capabilities. Dorminey et al. suggested a model called "MICE" [39]. This model replaces the factor of incentives/pressure through MICE (Money, Ideology, Coercion, and Ego) that represents the motivation of the individual to commit fraud. The "New Fraud Triangle" suggested by Kassem and Higson tried to combine all those more recent theories together into one model [78]. This model contains four elements: Motivation (MICE model), opportunity, fraudsters'

capability (fraud diamond model), and (4) personal Integrity (fraud scale model).

An overview of articles mentioned in this section can be found in Table 13. The articles are sorted based on appearance in the manuscript.

Table 13: Overview of the articles mentioned in Section 3.3.2: Psychological perspective.

Paper	**Key points**
Darley and Batson (1973)	Situational factor in explaining (un)ethical behavior
Walker and Pitts (1998)	Individual factor in explaining (un)ethical behavior
Trevino (1986)	The interactionist model
Ferrel and Gresham (1985)	The integrated contingency framework
Jones (1991)	Add moral issue to Ferrel's framework
Shariff and Norenzayan (2011), Asch (1951), Milgram (1963), Darley and Batson (1973), Doris (1998)	Valid the integrated contingency framework in various experiments
Piaget (1932), Kohlberg and Hersh (1977)	The moral development model, including 6 stages and 3 levels
Rest et al. (2000)	Add 3 schemas to the moral development model
Bandura (1989)	The social cognitive theory
Mead et al. (2009), Gino et al. (2011)	Experiments about relationship between self-control and unethical behavior
Barnes et al. (2011)	A lack of sleep could diminish self-control
Bandura (1999)	Self-sanctions
Aquino et al. (2009)	Moral identity
Mazar, Amir, and Ariely (2008)	People would like to maintain self-consistency
Detert, Treviño, and Sweitzer (2008)	Moral disengagement deactivates self-regulation
Gino and Pierce (2009), Vohs, Mead, and Goode (2006)	Money plays an important role in decision-making
Gino, Ayal, and Ariely (2009, 2013)	People might behave unethically to benefit their group rather than self-interests
Thau et al. (2015)	Employees' pro-group unethical behavior
Cressey (1950)	The fraud triangle framework
Ramamoorti (2008)	Incentives and perceived pressure lead to fraud
Batson et al. (1997), Tenbrunsel and Messick (2004), Gino and Ariely (2012), Gneezy (2005), Shalvi et al. (2015)	Various researches about rationalization
Festinger (1962)	Cognitive dissonance
Albrecht, Howe, and Romney (1984)	Fraud scale model
Wolfe and Hermanson (2004)	Fraud diamond model
Dorminey et al. (2012)	The MICE model
Kassem and Higson (2012)	New fraud triangle model

3.3.3 Behavior economic perspective

Behavior economics, in addition to the standard economics, also includes insights from psychology for understanding fraudulent behavior [90]. Simon postulated that people might have a bounded rationality instead of being fully rational [125]. He noticed that if we want to know why people act like they do, in addition to the economical consideration, the complexity of the environment and the limitations of the decision-maker itself have to be considered as well.

Tversky and Kahneman criticized the standard economic view including the expected utility theory and presented the prospect theory. This theory integrates psychological mechanisms like risk aversion/seeking in choices that involve gains/losses to predict behavior under uncertainty [135, 75, 76].

Illustrative examples from the field of behavior economics, showing that not only economical external rewards are involved when people decide to act (un)ethically, come from Ernst Fehr [50, 46, 49, 47, 45, 63]. Fehr and colleagues have demonstrated in several experiments such as the ultimatum game, dictator game and public good game that many people are not only maximizing their own profit but also concerning about social comparison, fairness, and the desire to reciprocate. For instance, in a one-time-only ultimatum game, the proposer offers the receiver a division of the money and the receiver then has to decide whether he accepts the offer or not. If the receiver rejects the proposed offer, both players go home without any money. From the standard economic perspective, the proposer should favor himself and split the money unequally and the receiver would even accept the smallest amount possible, because it is still more than nothing. However, in reality a majority of the offers are split equally, and many divisions that are not divided equally are rejected by the receivers. These results show that individuals consider social utility and others' outcomes rather than simply maximizing self-interested. Henrich et al. investigated the behavior of individuals from 15 societies in 12 countries with the ultimatum game [68]. The sample consisted of small-scale agriculturalist societies and nomadic tribes. Two main results from this study support the hypothesis of internalized reward mechanisms. First, the observed behavior varied fundamentally across societies. Second, the

individuals' preferences were not exogenous, as the standard economic model would predict, but rather shaped by their society's daily economic and their social interactions [90].

An overview of articles mentioned in this section can be found in Table 14. The articles are sorted based on appearance in the manuscript.

Table 14: Overview of the articles mentioned in Section 3.3.3: Behavior economic perspective.

Paper	**Key points**
Simon (1955)	Bounded rationality
Tversky and Kahneman (1975), Kahneman and Tversky (1979, 1984)	Prospect theory
Fehr and Schmidt (1999, 2001), Fehr and Gächter (2000, 2002), Fehr, Fischbacher, and Gächter (2002), Gintis et al. (2008)	Ultimatum game, dictator game, and public good game: people are not only maximizing their own profit but also caring about social comparison, fairness, and the desire to reciprocate
Henrich et al. (2001)	Individuals' preferences are not exogenous but rather shaped by their society's daily economic and their social interactions

3.3.4 Neuroscientific perspective

Recent findings coming from the field of neuroscience give promising insights for further evidence for the existence of internalized reward mechanisms. Neuroscience measures spatial and temporal brain activity with the aim to get a better understanding of how the brain is working [24]. Through measuring brain activities, brain functions are inferred. Neuroscience has made groundbreaking developments due to the rapid diffusion of brain imaging studies such as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) [90].

Research about fraudulent behavior or behavior in general focused for a long time on moral reasoning. But later evidence suggested that moral judgment is more about emotion and intuition than deliberate reasoning [67]. In neuroscience a substantial discussion aroused about whether unethical behavior emerges by deliberate, controlled or automatic, intuitive cognitive processes. The two competing theories are called "will" and "grace" hypothesis [66]. The "will" hypothesis states that ethical behavior results from the active exercise of self-control, or active

resistance of temptation. The “grace” hypothesis, on the other hand, states that honesty happens in the absence of temptation, or more automatically and intuitively [1, 53, 142]. Abe and Greene suggested reconciling the two different theories by demonstrating that the drivers of ethical behavior depend on the desirability of the reward. For ”grace” hypothesis, they conclude that honesty flows automatically in presence of weak neural responses to anticipated rewards. For ”will” one needs to refrain from dishonest behaviors in cases of relative high neural responses to anticipated rewards [1]. Experiments also showed that the cognitive control is mainly processed by the dorsolateral prefrontal cortex [48].

An overview of articles mentioned in this section can be found in Table 15. The articles are sorted based on appearance in the manuscript.

Table 15: Overview of the articles mentioned in Section 3.3.4: Neuroscientific perspective.

Paper	**Key points**
Camerer, Loewenstein, and Prelec (2005)	Neuroscience measures spatial and temporal brain activity with the aim to get a better understanding of how brain works
Greene and Haidt (2002)	Moral judgment is more about emotion and intuition than deliberate reasoning
Greene and Paxton (2010), Abe and Greene (2014), Fioretti and Marden (2015), Xu and Ma (2015)	“Will” and “grace” hypothesis. “Will” states that ethical behavior results from the active exercise of self-control, or active resistance of temptation. “Grace” states that honesty happens in the absence of temptation, or more automatically and intuitively.
Fehr and Rangel (2011)	The cognitive control is mainly processed by the dorsolateral prefrontal cortex

3.4 Studies on ordinary unethical behavior

Early ethics researches focused more on what behavior is desirable by society and how people should behave to fulfill ethical guidelines. Recently studies have shifted to a more descriptive approach, where researchers try to understand why individuals cheat [121]. In this section we will focus on the intentional and unintentional unethical behavior committed by ordinary

people.

3.4.1 Why ordinary people engage in intentional unethical behavior

The research stream of ordinary people committing intentional unethical behavior generally focuses on behaviors that people know to be wrong but still engage in [55]. In the other words, researchers are interested in why and how do people rationalize their behavior.

Some studies suggested that people try to strive to enhance a positive self-concept and to behave unethically only to a certain extend so that they can profit from wrongdoing while still feel moral [88, 96, 121]. The experiments conducted by Mazar, Amir, and Ariely indicated that people behave dishonestly enough to profit, but honestly enough to delude themselves of their own integrity [89].

The framework of Shalvi et al. is an extension of the "self-concept maintenance" theory of Mazar, Amir, and Ariely [89, 121]. The framework distinguishes between anticipated and experienced dissonance to identify different self-justification mechanisms that emerge either before or after an unethical act is committed. Self-justification, or self-serving justification, helps people to overcome the experienced or anticipated gap between their desire to profit by behaving unethically and their view of themselves as being moral. In this framework the justification that emerges before an unethical act is called pre-violation justification route. The authors have identified three pre-violation justifications and three post-violation ones. The details of these terms are explained below.

The first pre-violation justification is ambiguity. Situations where the norms or rules are ambiguous are prone to pre-violation justification [15]. As demonstration for this mechanism a dice-rolling experiment was conducted where the participants got more money for higher numbers on the dice. In the experiment only the one who rolled the dice saw the outcome, this eliminated the threat of being caught and made cheating legitimate to the individuals. When people rolled the dice three times, they reported higher numbers for the first roll than when they rolled the dice only once. This phenomenon emerged, because rolling the dice only once needs lying by inventing a number that had not been observed. But when the dice was rolled

three times instead of only once, it allowed participants to report a higher number for the first roll because they maybe consecutively saw a higher number in the second and third roll and transferred this number. Inventing facts is a clear violation to the individual whereas mixing up or transfer facts are more ambiguous and easier to justify up-front [120]. It is also important to mention that creative people are better in inventing facts, which allows them to use ambiguous situations in a self-serving manner, even when they only observe one roll [54]. The self-serving altruism is the second pre-violation justification to cheat. It can be applied when a lie causes no harm to another person but benefits others. Furthermore, if the number of people who benefit from one's unethical behavior increase, altruistic cheating increases as well [57, 1]. Another way how people can justify their unethical behaviors before committing them is by having something like a mental account for recent pro-social behavior. Moral licensing, which is the third pre-violation justification, works like a moral balance scale. If an individual recently did a lot of moral actions then subsequent unethical behavior is more easily to justify.

The post-violation justifications are cleaning, confessing, and distancing. Cleansing is an act where an individual is trying to liberate himself. The liberated act can be physical (e.g. pain or religious fasting) or symbolic (e.g. washing hands). Confessing, either to a higher entity or another person, helps people to reduce the feeling of dissonance. If an individual cannot clean, deny or confess for the unethical behavior, they distance themselves from their unethical act by pointing to others' unethical deeds and using stricter criteria when judging others' unethical behavior.

To sum up the above-mentioned researches of ordinary people committing intentional unethical behavior, mechanisms of rationalization, together with the pre- and post-justifications, have bridged the inconsistencies between individual's desire to behave ethical and the actual immoral behavior. These mechanisms can help people to avoid the feelings of anxiety, guilt or other negative emotions, or to protect one's self-concept. It involves the justification of an unacceptable behavior, thought or feeling in a logical manner to avoid the true reason for the action [106]. Furthermore, these mechanisms prove that morality can be stretched individually [55].

An overview of articles mentioned in this section can be found in Table 16. The articles are sorted based on appearance in the manuscript.

Table 16: Overview of the articles mentioned in Section 3.4.1: Why ordinary people engage in intentional unethical behavior?

Paper	**Key points**
Markus and Wurf (1987), Monin and Jordan (2009)	People try to strive to enhance a positive self-concept and to behave unethically only to a certain extend so that they can profit from wrongdoing while still feel moral
Mazar, Amir, and Ariely (2008)	Self-concept maintenance theory: People behave dishonesty enough to profit, but honestly enough to delude themselves of their own integrity
Shalvi et al. (2015), Batson et al. (1997), Shalvi, Elder, and Bereby-Meyer (2012), Gino and Ariely (2012), Gino, Ayal, and Ariely (2013), Abe et al. (2014), Gino (2015)	Self-justification theory, including three pre-violation justifications and three post-violation justifications: Ambiguity, self-serving altruism, moral licensing, cleaning, confessing, and distancing
Tenbrunsel and Messick (2004)	In intentional unethical behavior, people know what they do is wrong and try to rationalize their wrongdoing by self-justification to reduce ethical dissonance

3.4.2 Why ordinary people engage in unintentional unethical behavior?

As mentioned in the previous section, when individuals engage in unethical behavior, they do it normally with intent. In this case, they know what they do is wrong and try to rationalize their wrongdoing by self-justification mechanisms to reduce ethical dissonance [131]. However, sometimes people do behave unethically without even noticing that they crossed ethical borders [55, 119]. This phenomenon is called bounded ethicality, a concept derived from the bounded rationality of Herbert Simon [125, 30, 79]. People are exposed to systematic and predictable ethical blind spots, where they do not recognize the ethical dimension of their decisions [119]. Unintentional unethical behavior cannot be captured fully by the "new fraud triangle" from Kassem and Higson due to the fact that the reasons or motivations why people behave un-

ethically is not always evident when people act unintentional [78]. Therefore, new framework is indeed needed to explain this type of unethical behavior.

Sezer, Gino, and Bazerman identified three sources of unethical blind spots: Implicit biases, temporal distance from an ethical dilemma, and decision biases that lead people to disregard and misevaluate others' ethical lapses [119].

Implicit bias is a cognitive process and explains how unconscious attitudes can lead an individual to act against his moral values. Some of this implicit biases are in-group favoritism, illusion of objectivity, and the fairness of judgments are egocentric [93, 131]. For instance, many elite U.S. universities favor so-called "legacy" students without being aware of it. "Legacy" students are the children of alumni. This practice might prevent other ambitious students, also those who are more qualified, from being admitted to these universities [17].

Another source of unintentional unethical behavior could be the temporal distance form moral decisions. Individuals suffer from temporal inconsistencies. They tend to overestimate the amount to which they will behave ethically in the future. Such forecast errors can be explained by the two opposing systems: The "want self" that wants immediate gratification, and the "should self" that wants to make moral, long-sighted and responsible decisions [94]. Before an ethical decision, an individual thinks it would behave in accordance with their "should self" and this, in turn, would coincide with his or her moral self-image [89]. However, the "want self" becomes dominant as soon as it is time to make the decision. The shortsighted profit becomes much more salient while the longsighted, deliberative ethical choice fades away as closer one comes to the decision [131]. After decision, when they evaluate their actions, individuals try to reduce the dissonance they feel from the two opposing systems [96]. Taken in aggregate, these findings suggest that temporal inconsistencies prevent us from being as morally as we actually want to be.

The last source identified is the decision biases that lead people to ignore the immoral behaviors of others. Sezer, Gino, and Bazerman suggested several psychological mechanisms that influence the amounts to which people ignore other's immoral behaviors: The self-serving biases, the outcome bias, the presence of intermediaries, the gradual erosion of ethical behavior,

and no specific, identifiable victims. Self-serving bias means that an individual is less likely to notice, e.g. corrupt behavior of a co-worker if he can also benefit from the situation. The outcome bias describes the phenomenon that we judge exactly the same behavior as more ethical and worthy if it leads to a good outcome rather than a bad one. Furthermore, immoral behaviors of others are judged less harshly if the fraudster influences another person, i.e. the intermediary, to carry out the decision. According to a phenomenon known as the slippery slope effect, implicit biases prevent individuals from seeing gradual changes in their environment, including the gradual deterioration of ethical behavior [58, 29]. In addition, people tend to judge unethical behavior far more harshly when it harms specific, identifiable victims than when it harms a more anonymous group of people [62]. Together, these studies suggest that individuals ignore others' unethical behavior due to factors that have no particular relevance to the behavior's ethical content.

To sum up, the unconscious attitudes, the temporal inconsistencies between "want self" and "should self", and the ignoring of other's unethical behavior are the sources of unethical blind spots. With these blind spots, individuals are often not able to recognize the moral dimensions involved in their decision-making processes and the judgments they make about the behaviors of others, and are likely to commit unintentional unethical behavior.

An overview of articles mentioned in this section can be found in Table 17. The articles are sorted based on appearance in the manuscript.

3.5 Overview of statistical fraud detection techniques

In previous sections, we review the most important theories and concepts in unethical behavior. Fraud, defined as wrongful or criminal deception intended to result in financial or personal gain, is one form of unethical behavior. However, fraud is different from ordinary unethical behavior as it is about intentional use of deceits, tricks or some dishonest means to deprive another's money, property or legal right. Because of this, people have developed various methods to counter fraud.

Fraud prevention uses different measures to stop fraud from happening in the beginning. It

Table 17: Overview of the articles mentioned in Section 3.4.2: Why ordinary people engage in unintentional unethical behavior?

Paper	**Key points**
Chugh, Bazerman, and Banaji (2005), Kern and Chugh (2009), Simon (1955)	Bounded ethicality, which is a concept derived from the bounded rationality of Herbert Simon
Sezer, Gino, and Bazerman (2015), Gino (2015)	Three sources of unethical blind spots to explain unintentional unethical behavior: Implicit biases, temporal distance from an ethical dilemma, and decision biases that lead people to disregard and misevaluate others' ethical lapses
Messick and Bazerman (1996), Tenbrunsel and Messick (2004), Bazerman and Tenbrunsel (2011)	Implicit biases
Metcalfe and Mischel (1999), Mazar, Amir, and Ariely (2008), Tenbrunsel and Messick (2004), Monin and Jordan (2009)	Temporal distance from moral decision
Sezer, Gino, and Bazerman (2015), Gino and Bazerman (2009), Chugh (2004), Gino, Shu, and Bazerman (2010)	Decision biases that lead people to ignore the immoral behaviors of others: The self-serving biases, the outcome biases, the presence of intermediaries, the gradual erosion of ethical behavior, and no specific, identifiable victims

focuses on identifying and stopping existing fraud. Its key measures apply in creating a culture of honesty among people and organizations, create effective organization to minimize the risk of fraud, eliminate any fraud opportunities and create a comprehensive approach to fighting fraud.

Fraud deterrence deals with the causal factors of fraud. It is based on the assumption that fraud does not occur randomly but rather happens when some circumstances are present. The purpose of fraud deterrence is to fight against the root causes of fraud and what allows fraud to occur. The distinction between fraud prevention and deterrence is that the former involves identifying and stopping existing fraud whereas the latter focuses on eliminating factors that may cause fraud.

Fraud detection acts if fraud prevention and fraud deterrence have failed. It is defined by

diagnosing fraud in the fastest possible way once occurred. Obviously, the process of fraud detection must always be carried on since the failure to prevent and deter fraud is not an exact science with a definite result. It's also an endless evolving process: No matter how efficient one method might be to detect fraud, fraudsters will continuously find failures in the existing system. Indeed, criminals constantly adapt their strategies depending on the level and number of security barriers they have to cross. The main issue in developing new detection techniques is that the transfer of ideas is scarce since it would allow the fraudsters to get access more easily to valuable information to pass through the detection systems. Moreover, relevant real data sets are usually really hard to get access to in order to test new methods. Most of the time, companies prefer protecting their image rather than giving information about their exposure to fraud. Since new fraudsters come every day, all the existing detection techniques, from ancestor methods to the latest technologies, have to be applied since fraudsters might not be aware of all of them [21].

Statistical fraud detection techniques are varied, but the main idea is always to compare observed data with expected values. Depending on how the expected values are derived, there are two main domains in fraud detection: supervised and unsupervised methods.

In the supervised domain, both fraudulent and non-fraudulent data samples are employed to build detection models. It requires having both types of data and being sure about their affiliation. Those techniques are efficient only in detecting frauds that already happened in the past.

Traditional statistical classification methods, such as linear discriminant analysis and logistic discrimination, have been proved to be effective tools for many applications [91]. The idea of these methods is to find a linear combination of features that characterizes or separates two or more objects or events. For instance, Chae et al. employed logistic regression model to detect phantom transaction [26]. Phantom transaction is a kind fraud happening in online auction due to the collusion of sellers and buyers. The auction is falsely reported as "completed" while no item is actually exchanged. The creditor receive the money from the website whereas the credit card company delay the settlement with the online auction because no money is received

from the debtor. The authors showed that the use of "starting bid", "auction length", and "seller credit" in a logistic regression model could be very helpful for detection of phantom transaction.

Neural network is an artificial intelligence algorithm. Although it's sometimes referred to as "black box", in the sense that while it can approximate any function, studying its structure won't give you any insights on the structure of the function being approximated, it is a more powerful tool than linear discriminant analysis. For instance, Altman, Marco, and Varetto compared the performance of linear discriminant with that of neural network empirically [8]. Their comparisons showed that: Neural networks are able to approximate the numeric values of the scores generated by the discriminant functions; besides, neural networks are able to accurately classify groups of businesses as to their financial and operating health, with results that are very close to or, in some cases, even better than those of the discriminant analysis. Neural networks are commonly used in telecommunication fraud due to their capacity for representing complex and non-linear models and not having severe limitations and assumptions concerning the type of input data. In subscription fraud related to fixed telephone lines, neural network have been used as classifiers or predictive systems. As a classifier, a feed forward neural network was used to classify subscribers as fraudulent [69]. As a predictive system, a multilayer feed forward perceptron neural network was used to predict whether a new phone line correspond to fraud [41]. Regarding mobile phone detection, a bidirectional artificial neural network was able to predict fraud comparing two unidirectional Artificial Neural Networks using time-series representing individuals' behavior [84].

Rule-based methods are algorithms that produce classifiers using various rules. For instance, Clark and Niblett designed the CN2 induction system, which can effectively induct simple, comprehensible rules in domains where problems of poor description language and/or noise may be present [31]. A framework was proposed to detect superimposed fraud based on rules [44]. Using this technique, indicators of fraudulent behavior are uncovered by using adaptive rules in order to build monitors. Then those monitors are employed to give a description of fraudulent behavior profiles and serve as input in the detection model. The detector constructor network integrates all the different monitors generated through the user's data on a daily basis

and then detects a fraudulent behavior.

Tree-based methods have similar forms as rule-based ones. Decision trees are table of trees shape with connecting lines to available nodes. Each node is either connected with more nodes or a leaf node which defines a classification. There are two main types of decision trees. Classification tree analysis is when the predicted outcome is the class to which the data belongs. Regression tree analysis is when the predicted outcome can be considered a real number. Those classifiers, such as CHAID [77], CART [23], C4.5 [108] can work alone, in parallel or can be combined together in order to create better detection performances. The combination of several classifiers creates meta-classifiers and increases the detection accuracy [27].

On the other hand, unsupervised methods are used when there are no previous fraudulent and legitimate observations that can be classified. The idea here is to model a baseline that represent legitimate referential behavior and then try to detect observations that are significant deviated from this referential. These observations are named as anomalies and outliers. The advantage in unsupervised methods is to be able to identify frauds that were not discovered before.

Peer-group analysis has been developed using behavioral fraud detection. It is a method that enables the identification of accounts that behave in a different manner than others at a certain point in time while they had the same behavioral previously [22, 72]. While peer-group analysis implies a behavioral change of a certain account from others, break-point analysis uses different transactions from a single card user. If a behavioral change of a single card is identified, such as a sudden transaction of a high amount or a high frequency usage, then it is flagged as "suspicious" [22]. Hidden Markov Model (HMM) is a semi-supervised anomaly detection technique. It is a stochastic model with a finite set of states. Each state is related with a probability distribution. Transitions among those states are given under a set of probabilities. Each state can be associated with an outcome depending on the probability distribution. The HMM is initially trained with the normal behavior of a cardholder. Then, the cardholder's profile is assigned as low, normal and high spending based on their previous spending behavior. Each cardholder is assigned with a set of probability regarding the amount of transactions. Any

new transaction's amount is affiliated within a category and compared to a defined threshold in order to distinguish between fraudulent and legitimate transactions [129].

In the end, it's worth to mention that various fraud detection techniques can be combined together. For instance, a hybrid forecasting system was created to detect fraudulent financial statements across firms [83]. The system is composed by decision trees, artificial neural networks, Bayesian networks, rules-learners and support vector machines. The stacking techniques consist of combining all those methods together to increase the efficiency. An extreme outlier elimination and hybrid sampling technique were proposed to counter skewed datasets in insurance claims [103]. K-nearest neighbors' algorithm was used to detect extreme outliers and a hybrid sampling was implemented to improve the detection accuracy. Panigrahi et al. proposed a detection technique which includes the rule-based filtering, Dempster-Shafer theory, and Bayesian learning [105]. Evidences from on rule-based filtering are combined and associated by using Dempster-Shafer theory in order to calculate a primary belief on each transaction. Afterwards, Bayesian learning, which is a statistical tool updating previous behavioral pattern to help the classification of a new transaction, provides a suspicion score to classify the transaction.

An overview of articles mentioned in this section can be found in Table 18. The articles are sorted based on appearance in the manuscript.

3.6 Network science in fraud detection

Social networks analysis is a useful tool in fraud detection models. Observations or individuals are usually represented as nodes, and the relationships among them are edges. There are various tools which employ social network analysis in fraud detection. For instance, link analysis tries to relate known fraudsters to other individuals by using record linkage [139]. In telecommunication fraud, fraudsters are found to often call the same numbers from another account once the previous account has been disabled for fraud [33]. Oddball algorithm, which was proposed by Akoglu, McGlohon, and Faloutsos, can be used to detect anomalous nodes in (un)weighted graphs [4]. The authors noticed that "egonets", the induced sub- graph of the node of interest

Table 18: Overview of the articles mentioned in Section 3.5: Overview of statistical fraud detection techniques.

Supervised methods	
Paper	**Key points**
Bolton and Hand (2002)	A review of statistical fraud detection
Geoffrey (1992), Chae et al. (2007)	Traditional statistical classification methods, e.g. linear discriminant analysis and logistic discrimination
Altman, Marco, and Varetto (1994), Hilas and Mastorocostas (2008), Estvez, Held, and Perez (2006), Krenker et al. (2009)	Use neural network as classifiers or predictive systems in fraud detection
Clark and Niblett (1989), Fawcett and Provost (1997)	Rule-based methods as classifiers
Kass (1980), Breiman et al. (1984), Quinlan (1993), Chan et al. (1999)	Tree-based methods as classifiers
Unsupervised methods	
Paper	**Key points**
Bolton and Hand (2001), Hui et al. (2014)	Peer-group analysis as anomaly detection techniques
Srivastava et al. (2008)	Hidden Markov Model as semi-supervised anomaly detection technique
Hybrid methods	
Paper	**Key points**
Kotsiantis et al. (2016)	Hybrid forecasting system, including decision trees, artificial neural networks, Bayesian networks, rules-learners, and support vector machines
Padmaja et al. (2007)	K-nearest neighbors' algorithm is used to detect extreme outliers and a hybrid sampling is implemented to improve the detection accuracy
Panigrahi et al. (2009)	Use rule-based filtering, Dempster-Shafer theory, and Bayesian learning to classify new transactions with suspicion scores.

and its neighbors, follow certain patterns in density, weights, principle eigenvalues, and ranks. They then employed "egonets" to detect anomalies.

Another worth-mentioning technique is NetProbe. NetProbe algorithm is an online auction fraud detection system [104]. It shapes the auction website as a network composed by nodes representing buyers and sellers and edges representing transactions. The concept is to deduce properties of a single user by examining properties of other related users. Given this graph, the likelihood of being a fraudster is calculated using the user's immediate neighbors. This model uses Markov random field to detect suspicious pattern in the network and then uses belief propagation algorithm to localize fraudulent trades. Belief propagation algorithm is used to perform inference on graphs. An incremental version of NetProbe was also created to quickly update beliefs when the graph topology changes. In this version, new edge does not alter the entire graph but changes the immediate neighborhood of an edge.

Community-based anomaly detection techniques have drawn a lot of attentions recently. These approaches aim to find the graph objects, e.g. nodes, edges, substructures, and the patterns of interactions that are rare and differ significantly from the majority of the reference objects in the graph [5, 28, 117]. Varies community detection methods have been developed to discover node partitions and can be applied to monitor the structural or contextual change in every community [64, 32, 99, 112, 116, 109, 20, 87]. The descriptions of these methods as well as the comparisons of their performances can be found in Chapter 4.

An overview of articles mentioned in this section can be found in Table 19. The articles are sorted based on appearance in the manuscript.

3.7 Managerial implications of studies on unethical behavior

In this section we provide some practical advices for marketing managers on how to curb intentional and unintentional unethical behaviors of their employees/costumers.

3.7.1 How to reduce intentional unethical behavior?

Before a manager starts to implement an action to curb cheating behavior, he should try to find out whether the cheating behavior of the employees/costumers is caused by external rewards or

Table 19: Overview of the articles mentioned in Section 3.6: Network science in fraud detection.

Paper	**Key points**
Wasserman and Faust (1994), Cortes, Pregibon, and Volinsky (2001)	Use link analysis to relate known fraudsters to the other individuals
Akoglu, McGlohon, and Faloutsos (2010)	Oddball algorithm, which employs "egonets" to detect anomalies
Pandit, Wang, and Faloutsos (2007)	NetProbe, which uses Markov random field to detect suspicious pattern in the network, and belief propagation algorithm to localize fraudulent trades
Akoglu, Tong, and Koutra (2015), Girvan and Newman (2002), Clauset, Newman, and Moore (2004), Newman (2006), Reichardt and Bornholdt (2006), Rosvall and Bergstrom (2007), Raghavan, Albert, and Kumara (2007), Blondel et al. (2008), Leskovec, Lang, and Mahoney (2010), Savage et al. (2014), Chen, Hendrix, and Samatova (2012)	Various community detection algorithms can be used in the community-based anomaly detection techniques in order to find the graph objects, e.g. nodes, edges, substructures, and the patterns of interactions that are rare and differ significantly from the majority of the reference objects in the graph

internal individual reward and/or punishment mechanisms [90]. If a manager ignores this first step, he is likely to choose the wrong tools to curb cheating behavior and hence, effectiveness would suffer.

The managerial intervention is simple, if the reason of unethical behavior lies only in the greater external rewards than the costs of cheating. In that purely economic case, the manager's task should be to make the costs of unethical behavior larger than the expected rewards. This can be attained by increasing the probability of being caught or by increasing the magnitude of the punishment. The probability of being caught is much more effective than the severity of the punishment [19].

If unethical behavior is caused by internal individual reward and/or punishment mechanisms, then it gets more complicated and many tools can be applied depending on what's the internal reason of cheating. Firstly, a lack of internalized social norms could be the cause for

unethical behavior [25]. In this case a manager should set up sessions or meetings to educate the employees. The employees can internalize desirable social norms and rules, and consecutively strengthen the internal reward mechanisms. Internalizing norms needs its time, and hence, is a long-run tool to curb cheating behavior [90]. Experiments from the area of insurance fraud claim that fraud detection systems have their limitations for customers who perceive a low detection probability [43]. For such costumers, the limited success of fraud detection systems reinforces their intentions towards engaging in insurance fraud. They do not fear of being caught but rather have internal mechanisms of justification. Therefore, it is also important to establish moral norms, e.g. inform costumers about the negative societal consequences of insurance fraud or state through advertisement that insurance fraud is a highly unethical act, rather than simply increase the probability of being caught. Secondly, a short-run tool to reduce cheating behavior is used if the cause of unethical behavior is a low moral identity (or lack of self-awareness). The moral identity influences the way the internalized social norms are activated [12]. Therefore, a manager should make use of contextual cues that increases the moral identity and boost the activation of the internalized norms. One way to direct attention towards the moral identity is to let the employees sign an "honor code" or a "code of conduct" [90, 123]. Another effective way to activate the moral identity is to let people sign their names at the beginning of a paper (e.g. an employment contract) rather than the conventional way of signing at the very end. This increases moral identity right before it is needed most, which in turn, promotes moral behavior [124]. Thirdly, if the unethical behavior is caused by self-deception, it is hard for a manager to curb unethical behavior. Self-deception based on a self-serving bias is very stable and cannot be changed by specific training or education. Therefore, the most effective way to reduce self-deception is to eliminate the situations that lead to such biases. Such incentives/situations can emerge if a manager does not establish clear and strict rules which have to be followed in the organization [90].

Management courses usually teach that the performance of employees increases more when specific goals are set rather than vague ones. Goal setting has become an important role in managing employees' motivation. Schweitzer, Ordóñez, and Douma conducted several experiments

to study the behavior of overstating productivity [118]. They figured out that participant with unmet goals did more often overstate productivity than participants who were told to do their best. Furthermore, participants with unmet reward goals did more often overstate productivity than participants with unmet mere goals. Their findings suggest that managers should set goals carefully and especially set goals that are achievable for employees.

The last point to consider is the self-depletion issue. Self-control depletion is a source that promotes unethical behavior [61]. When self-control resources are depleted, individuals don't have enough cognitive resources to recognize the moral dimension in the decision-making process and hence, temptations to behave in an unethical way rise. Gino and colleagues suggested that managers should try to remove temptations, to develop self-control, and to monitor individuals that are prone to be depleted fast (e.g. employees who are often interrupted at work).

Experiments conducted by Barnes et al. proved that a lack of sleep could diminish self-control, which in turn reduces the inhibition of unethical behavior [14]. Implications for managers are to create an organizational climate that is less stressful so that employees have more and a qualitative better rest.

3.7.2 How to reduce unintentional unethical behavior?

Reducing the unintentional unethical behavior is a complex and difficult task because the perpetrators themselves are unaware that they cross moral boundaries. Interventions that address intentional unethical behavior are not always effective for unintentional unethical behavior as well [144]. Sezer and colleagues presented two categories that help individuals to overcome the underlying biases that lead them to commit unethical behavior: Moving from system 1 to system 2 thinking, and strategies aimed at institutional design [119].

There exist two thinking modes in humans that can overrule each other depending on the situation [74]. System 1 thinking is fast, automatic, emotional, intuitive and doesn't need much effort, whereas system 2 thinking is slow, deliberate, reason based, rational and needs much more effort. Research has shown that system 1 is much more biased than system 2. If individuals are exhausted or tired, they rely on system 1. As system 1 is more biased, individuals

are prone to commit unethical behavior [92, 61, 119]. Therefore, the managerial goal should be to shift individuals thinking from system 1 to system 2. Managers' interventions to achieve this goal can be either changing the framing of the decision or giving individuals more time for deliberate thinking, respectively taking away time pressure. Such interventions can reduce automatic, undeliberate decision-making and curb unethical behavior to a certain extent [67, 120]. Kern and Chugh found evidence in three experiments that individuals are more likely to behave unethical if a decision is presented in a loss frame than if the decision is presented in a gain frame [79]. Another useful tool for managers is to present their employees options in a joint evaluation. Bazerman et al. argued that joint decision-making, which compares two or more options simultaneously, is less emotive and more reasoned than separate decision-making that thinks about one option at a time [16].

The second category mentioned above is "strategies aimed at institutional design". Strategies should be designed in a way that they can disclosure or eliminate conflicts of interests and promote people toward more ethical choices. For example, signing an "honor code" and not setting unrealistically high goals can change people's incentives and tasks, and thus prevent unethical behavior, while conflicts of interests stop auditors from making unbiased decisions about their own clients in the auditing industry [97].

3.8 Conclusion and future steps

In the project, various articles from four different areas have been reviewed in order to gain a deeper understanding of unethical behavior. Although from the standard economics perceptive committing fraud does not have any internal costs, latest evidences from psychology and behavior economics suggest that individuals do have internal costs while committing fraud. These costs emerge because the actions an individual undertakes, while engaging in cheating behavior, are inconsistent with the internal ethical standards and moral principles. A growing body of research points to the fact that not only there exist internal psychological costs while cheating, but also that the costs are different among individuals. Therefore, as every individual might have very different internal costs, there exists no universal formula to predict precisely whether

an individual commits fraud or not in a certain situation. Moreover, due to self-serving biases and bounded ethicality, people are sometimes not even aware of crossing ethical borders, and hence there are not many possibilities to reduce such unethical behavior. Further development in neuroscience might shed light on understanding the underlying mechanisms behind these internal psychological costs, and this will help us to prevent, detect, and punish fraudulent behavior more efficiently. Under the current situation, we believe that the "new fraud triangle" framework from Kassem and Higson [78] is a simple yet efficient tool to understand why individuals engage in unethical behavior. But it needs to integrate the mechanisms of ordinary unethical behavior, especially the ones of unintentional unethical behavior in order to be more complete.

We have also examined many different fraud detection techniques. The attempts to commit fraud have drastically increased over the years, which make the field of fraud detection more important than ever. Millions of dollars, euros and others currencies are lost every year as fraudsters keep on finding new failures in the existing systems. Different statistical fraud detection techniques, such as linear discriminant analysis, neural network, rule-based and tree-based classification methods, anomaly and outlier detections should all be used as fraudsters might not be aware of all of them. We would like to strengthen the fact that network science can be widely applied in fraud detection, as it is a good tool to detect anomalies in both individuals and the interactions between them.

Although some useful suggestions to reduce unethical behavior are proposed in the previous section, managers will never be able to erase cheating behavior fully. To quote the words by Bolton and Hand [21]: "Fraud can be reduced to as low a level one likes, but only by virtue of a corresponding level of effort and cost." Therefore, we would strongly encourage mangers to employ different methods in a proper way in order to best detect and reduce the possible unethical behaviors of their customers and employees.

References

[1] Abe, N. *et al.* The neural basis of dishonest decisions that serve to harm or help the target. *Brain and cognition* **90**, 41–49 (2014).

[2] Abe, N. & Greene, J. D. Response to anticipated reward in the nucleus accumbens predicts behavior in an independent test of honesty. *The Journal of Neuroscience* **34**, 10564–10572 (2014).

[3] ACFE. The Fraud Triangle (2016). URL `http://www.acfe.com/fraud-triangle.aspx`.

[4] Akoglu, L., McGlohon, M. & Faloutsos, C. Oddball: Spotting anomalies in weighted graphs. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, 410–421 (Springer, 2010).

[5] Akoglu, L., Tong, H. & Koutra, D. Graph based anomaly detection and description: a survey. *Data Mining and Knowledge Discovery* **29**, 626–688 (2015).

[6] Albrecht, W. S., Howe, K. R. & Romney, M. B. *Deterring fraud: the internal auditor's perspective* (Inst of Internal Auditors, 1984).

[7] Allingham, M. G. & Sandmo, A. Income tax evasion: A (1972).

[8] Altman, E. I., Marco, G. & Varetto, F. Corporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks (the italian experience). *Journal of banking & finance* **18**, 505–529 (1994).

[9] Aquino, K., Freeman, D., Reed II, A., Lim, V. K. & Felps, W. Testing a social-cognitive model of moral behavior: the interactive influence of situations and moral identity centrality. *Journal of personality and social psychology* **97**, 123 (2009).

[10] Asch, S. E. Effects of group pressure upon the modification and distortion of judgments. *Groups, leadership, and men* 222–236 (1951).

[11] Bandura, A. Human agency in social cognitive theory. *American psychologist* **44**, 1175 (1989).

[12] Bandura, A. Moral disengagement in the perpetration of inhumanities. *Personality and social psychology review* **3**, 193–209 (1999).

[13] Bandura, A. Social cognitive theory: An agentic perspective. *Annual review of psychology* **52**, 1–26 (2001).

[14] Barnes, C. M., Schaubroeck, J., Huth, M. & Ghumman, S. Lack of sleep and unethical conduct. *Organizational Behavior and Human Decision Processes* **115**, 169–180 (2011).

[15] Batson, C. D., Kobrynowicz, D., Dinnerstein, J. L., Kampf, H. C. & Wilson, A. D. In a very different voice: unmasking moral hypocrisy. *Journal of personality and social psychology* **72**, 1335 (1997).

[16] Bazerman, M. H., Gino, F., Shu, L. L. & Tsay, C.-J. Joint evaluation as a real-world tool for managing emotional assessments of morality. *Emotion Review* **3**, 290–292 (2011).

[17] Bazerman, M. H. & Tenbrunsel, A. E. *Blind spots: Why we fail to do what's right and what to do about it* (Princeton University Press, 2011).

[18] Becker, G. S. Irrational behavior and economic theory. *The Journal of Political Economy* **70**, 1–13 (1962).

[19] Becker, G. S. *Crime and Punishment: An Economic Approach*, vol. 76 (1968).

[20] Blondel, V. D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment* **2008**, P10008 (2008).

[21] Bolton, R. J. & Hand, D. J. Statistical fraud detection: A review. *Statistical science* 235–249 (2002).

[22] Bolton, R. J. & Hand, D. J. Peer group analysis–local anomaly detection in longitudinal data. Tech. Rep., Citeseer (2001).

[23] Breiman, L., Friedman, J., Stone, C. J. & Olshen, R. A. *Classification and regression trees* (CRC press, 1984).

[24] Camerer, C., Loewenstein, G. & Prelec, D. Neuroeconomics: How neuroscience can inform economics. *Journal of economic Literature* **43**, 9–64 (2005).

[25] Campbell, E. Q. The internalization of moral norms. *Sociometry* 391–412 (1964).

[26] Chae, M., Shim, S., Cho, H. & Lee, B. An empirical analysis of fraud detection in online auctions: Credit card phantom transaction. In *System Sciences, 2007. HICSS 2007. 40th Annual Hawaii International Conference on*, 155a–155a (IEEE, 2007).

[27] Chan, P. K., Fan, W., Prodromidis, A. L. & Stolfo, S. J. Distributed data mining in credit card fraud detection. *IEEE Intelligent Systems and Their Applications* **14**, 67–74 (1999).

[28] Chen, Z., Hendrix, W. & Samatova, N. F. Community-based anomaly detection in evolutionary networks. *Journal of Intelligent Information Systems* **39**, 59–85 (2012).

[29] Chugh, D. Societal and managerial implications of implicit social cognition: Why milliseconds matter. *Social Justice Research* **17**, 203–222 (2004).

[30] Chugh, D., Bazerman, M. H. & Banaji, M. R. Bounded ethicality as a psychological barrier to recognizing conflicts of interest. *Conflicts of interest: Challenges and solutions in business, law, medicine, and public policy* 74–95 (2005).

[31] Clark, P. & Niblett, T. The cn2 induction algorithm. *Machine learning* **3**, 261–283 (1989).

[32] Clauset, A., Newman, M. E. & Moore, C. Finding community structure in very large networks. *Physical review E* **70**, 066111 (2004).

[33] Cortes, C., Pregibon, D. & Volinsky, C. Communities of Interest. *International Conference on Advances in Intelligent Data Analysis* **2189**, 105–114 (2001).

[34] Cressey, D. R. The criminal violation of financial trust. *American Sociological Review* **15**, 738–743 (1950).

[35] Darley, J. M. & Batson, C. D. " from jerusalem to jericho": A study of situational and dispositional variables in helping behavior. *Journal of Personality and Social Psychology* **27**, 100 (1973).

[36] Derrig, R. A. Insurance fraud. *Journal of Risk and Insurance* **69**, 271–287 (2002).

[37] Detert, J. R., Treviño, L. K. & Sweitzer, V. L. Moral disengagement in ethical decision making: a study of antecedents and outcomes. *Journal of Applied Psychology* **93**, 374 (2008).

[38] Doris, J. M. Persons, situations, and virtue ethics. *Nous* **32**, 504–530 (1998).

[39] Dorminey, J., Fleming, A. S., Kranacher, M.-J. & Riley Jr, R. A. The evolution of fraud theory. *Issues in Accounting Education* **27**, 555–579 (2012).

[40] Ernst & Young. Global Fraud Survey 2016 Justifying unethical behavior and misconduct (2016). URL `http://www.ey.com/GL/en/Services/Assurance/Fraud-Investigation---Dispute-Services/EY-global-fraud-survey-2016-justifying-unethical-behavior-and-misconduct`.

[41] Estévez, P. a., Held, C. M. & Perez, C. a. Subscription fraud prevention in telecommunications using fuzzy rules and neural networks. *Expert Systems with Applications* **31**, 337–344 (2006).

[42] Ethical Systems. Cheating & Honesty (2016). URL `http://www.ethicalsystems.org/content/cheating-honesty`.

[43] Farashah, A. D. & Estelami, H. The interplay of external punishment and internal rewards: An exploratory study of insurance fraud. *Journal of Financial Services Marketing* **19**, 253–264 (2014).

[44] Fawcett, T. & Provost, F. Adaptive fraud detection. *Data mining and knowledge discovery* **1**, 291–316 (1997).

[45] Fehr, E., Fischbacher, U. & Gächter, S. Strong reciprocity, human cooperation, and the enforcement of social norms. *Human nature* **13**, 1–25 (2002).

[46] Fehr, E. & Gächter, S. Fairness and retaliation: The economics of reciprocity. *The journal of economic perspectives* **14**, 159–181 (2000).

[47] Fehr, E. & Gächter, S. Altruistic punishment in humans. *Nature* **415**, 137–140 (2002).

[48] Fehr, E. & Rangel, A. Neuroeconomic foundations of economic choice – recent advances. *The Journal of Economic Perspectives* **25**, 3–30 (2011).

[49] Fehr, E. & Schmidt, K. M. Theories of fairness and reciprocity-evidence and economic applications (2001).

[50] Fehr, E. & Schmidt, K. M. A theory of fairness, competition, and cooperation. *Quarterly journal of Economics* 817–868 (1999).

[51] Ferrell, O. C. & Gresham, L. G. A contingency framework for understanding ethical decision making in marketing. *The Journal of Marketing* 87–96 (1985).

[52] Festinger, L. *A theory of cognitive dissonance*, vol. 2 (Stanford university press, 1962).

[53] Fioretti, M. & Marden, S. Suboptimal dishonesty: Rationality in the absence of strategic behavior in honesty experiments. *The Journal of Neuroscience* **35**, 1817–1818 (2015).

[54] Gino, F. & Ariely, D. The dark side of creativity: original thinkers can be more dishonest. *Journal of personality and social psychology* **102**, 445 (2012).

[55] Gino, F. Understanding ordinary unethical behavior: why people who value morality act immorally. *Current opinion in behavioral sciences* **3**, 107–111 (2015).

[56] Gino, F., Ayal, S. & Ariely, D. Contagion and differentiation in unethical behavior the effect of one bad apple on the barrel. *Psychological science* **20**, 393–398 (2009).

[57] Gino, F., Ayal, S. & Ariely, D. Self-serving altruism? the lure of unethical actions that benefit others. *Journal of economic behavior & organization* **93**, 285–292 (2013).

[58] Gino, F. & Bazerman, M. H. When misconduct goes unnoticed: The acceptability of gradual erosion in others' unethical behavior. *Journal of experimental Social psychology* **45**, 708–719 (2009).

[59] Gino, F. & Pierce, L. Dishonesty in the name of equity. *Psychological Science* **20**, 1153–1160 (2009).

[60] Gino, F. & Pierce, L. The abundance effect: Unethical behavior in the presence of wealth. *Organizational Behavior and Human Decision Processes* **109**, 142–155 (2009).

[61] Gino, F., Schweitzer, M. E., Mead, N. L. & Ariely, D. Unable to resist temptation: How self-control depletion promotes unethical behavior. *Organizational Behavior and Human Decision Processes* **115**, 191–203 (2011).

[62] Gino, F., Shu, L. L. & Bazerman, M. H. Nameless+ harmless= blameless: When seemingly irrelevant factors influence judgment of (un) ethical behavior. *Organizational Behavior and Human Decision Processes* **111**, 93–101 (2010).

[63] Gintis, H., Henrich, J., Bowles, S., Boyd, R. & Fehr, E. Strong reciprocity and the roots of human morality. *Social Justice Research* **21**, 241–253 (2008).

[64] Girvan, M. & Newman, M. E. Community structure in social and biological networks. *Proceedings of the national academy of sciences* **99**, 7821–7826 (2002).

[65] Gneezy, U. Deception: The role of consequences. *The American Economic Review* **95**, 384–394 (2005).

[66] Greene, J. D. & Paxton, J. M. Patterns of neural activity associated with honest and dishonest moral decisions. *Proceedings of the National Academy of Sciences* **106**, 12506–12511 (2009).

[67] Greene, J. & Haidt, J. How (and where) does moral judgment work? *Trends in cognitive sciences* **6**, 517–523 (2002).

[68] Henrich, J. *et al.* In search of homo economicus: behavioral experiments in 15 small-scale societies. *The American Economic Review* **91**, 73–78 (2001).

[69] Hilas, C. S. & Mastorocostas, P. A. An application of supervised and unsupervised learning approaches to telecommunications fraud detection. *Knowledge-Based Systems* **21**, 721–726 (2008).

[70] Hirschi, T. & Gottfredson, M. R. Self-control theory. *Explaining criminals and crime* 81–96 (2001).

[71] Hobbes, T. Leviathan, ed. cb macpherson. In *London: Penguin [1651]* (1968).

[72] Chi Hui, F., Koneru, V. C., Mat Ali, N. & Harun, S. Implementing peer group analysis within a track and trace system to detect potential fraud (s). *International Journal of Supply Chain Management* **3** (2014).

[73] Jones, T. M. Ethical decision making by individuals in organizations: An issue-contingent model. *Academy of management review* **16**, 366–395 (1991).

[74] Kahneman, D. *Thinking, fast and slow* (Macmillan, 2011).

[75] Kahneman, D. & Tversky, A. Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the econometric society* 263–291 (1979).

[76] Kahneman, D. & Tversky, A. Choices, values, and frames. *American psychologist* **39**, 341 (1984).

[77] Kass, G. V. An exploratory technique for investigating large quantities of categorical data. *Applied statistics* 119–127 (1980).

[78] Kassem, R. & Higson, A. The new fraud triangle model. *Journal of Emerging Trends in Economics and Management Sciences* **3**, 191 (2012).

[79] Kern, M. C. & Chugh, D. Bounded ethicality the perils of loss framing. *Psychological Science* **20**, 378–384 (2009).

[80] Kleemans, E. R. Organized crime and the visible hand: A theoretical critique on the economic analysis of organized crime. *Criminology and Criminal Justice* **13**, 615–629 (2013).

[81] Kleven, H. J., Knudsen, M. B., Kreiner, C. T., Pedersen, S. & Saez, E. Unwilling or unable to cheat? evidence from a tax audit experiment in denmark. *Econometrica* **79**, 651–692 (2011).

[82] Kohlberg, L. & Hersh, R. H. Moral development: A review of the theory. *Theory into practice* **16**, 53–59 (1977).

[83] Kotsiantis, S., Koumanakos, E., Tzelepis, D. & Tampakas, V. Forecasting fraudulent financial statements using data mining. *International Journal of Computational Intelligence* **3**, 104–110 (2006).

[84] Krenker, A., Volk, M., Sedlar, U., Bešter, J. & Kos, A. Bidirectional artificial neural networks for mobile-phone fraud detection. *Etri Journal* **31**, 92–94 (2009).

[85] Kroll. Global Fraud Report (2015). URL `http://anticorruzione.eu/wp-content/uploads/2015/09/Kroll_Global_Fraud_Report_2015low-copia.pdf`.

[86] Leinbach-Reyhle, N. New Report Identifies US Retailers Lose $60 Billion a Year, Employee Theft Top Concern (2015). URL `http://www.forbes.com/sites/nicoleleinbachreyhle/2015/10/07/new-report-identifies-us-retailers-lose-60-billion-a-year-employee-theft-top-concern/`.

[87] Leskovec, J., Lang, K. J. & Mahoney, M. Empirical comparison of algorithms for network community detection. In *Proceedings of the 19th international conference on World wide web*, 631–640 (ACM, 2010).

[88] Markus, H. & Wurf, E. The dynamic self-concept: A social psychological perspective. *Annual review of psychology* **38**, 299–337 (1987).

[89] Mazar, N., Amir, O. & Ariely, D. The dishonesty of honest people: A theory of self-concept maintenance. *Journal of marketing research* **45**, 633–644 (2008).

[90] Mazar, N. & Ariely, D. Dishonesty in everyday life and its policy implications. *Journal of public policy & Marketing* **25**, 117–126 (2006).

[91] Geoffrey, J. M. Discriminant analysis and statistical pattern recognition (1992).

[92] Mead, N. L., Baumeister, R. F., Gino, F., Schweitzer, M. E. & Ariely, D. Too tired to tell the truth: Self-control resource depletion and dishonesty. *Journal of experimental social psychology* **45**, 594–597 (2009).

[93] Messick, D. M. & Bazerman, M. H. Ethical leadership and the psychology of decision making. *MIT Sloan Management Review* **37**, 9 (1996).

[94] Metcalfe, J. & Mischel, W. A hot/cool-system analysis of delay of gratification: dynamics of willpower. *Psychological review* **106**, 3 (1999).

[95] Milgram, S. Behavioral study of obedience. *The Journal of abnormal and social psychology* **67**, 371 (1963).

[96] Monin, B. & Jordan, A. H. The dynamic moral self: A social psychological perspective. *Personality, identity, and character: Explorations in moral psychology* 341–354 (2009).

[97] Moore, D. A., Tetlock, P. E., Tanlu, L. & Bazerman, M. H. Conflicts of interest and the case of auditor independence: Moral seduction and strategic issue cycling. *Academy of Management Review* **31**, 10–29 (2006).

[98] Nagin, D. S. & Pogarsky, G. An experimental investigation of deterrence: Cheating, self-serving bias, and impulsivity. *Criminology* **41**, 167–194 (2003).

[99] Newman, M. E. Finding community structure in networks using the eigenvectors of matrices. *Physical review E* **74**, 036104 (2006).

[100] Oppel Jr., R. A. & Sorkin, A. R. Enron's Collapse: The Overview; Enron Collapses as suitor cancels plans for merger (2001). URL `http://www.nytimes.com/2001/11/29/business/enron-s-collapse-the-overview-enron-collapses-as-suitor-cancels-plans-for-merger.html?pagewanted=all`.

[101] Oxford Dictionaries. Definition of fraud in English (2015). URL http://www.oxforddictionaries.com/definition/english/fraud.

[102] Oxford Dictionaries. Definition of cheat in English (2015). URL http://www.oxforddictionaries.com/definition/english/cheat?q=cheating.

[103] Padmaja, T. M., Dhulipalla, N., Bapi, R. S. & Krishna, P. R. Unbalanced data classification using extreme outlier elimination and sampling techniques for fraud detection. In *Advanced Computing and Communications, 2007. ADCOM 2007. International Conference on*, 511–516 (IEEE, 2007).

[104] Pandit, S., Chau, D. H., Wang, S. & Faloutsos, C. Netprobe: a fast and scalable system for fraud detection in online auction networks. In *Proceedings of the 16th international conference on World Wide Web*, 201–210 (ACM, 2007).

[105] Panigrahi, S., Kundu, A., Sural, S. & Majumdar, A. K. Credit card fraud detection: A fusion approach using dempster–shafer theory and bayesian learning. *Information Fusion* **10**, 354–363 (2009).

[106] Pedersen, T. Rationalization (2016). URL http://psychcentral.com/encyclopedia/rationalization/.

[107] Piaget, J. The moral development of the child. *Kegan Paul, London* (1932).

[108] Quinlan, J. R. C4. 5: Programming for machine learning. *Morgan Kauffmann* 38 (1993).

[109] Raghavan, U. N., Albert, R. & Kumara, S. Near linear time algorithm to detect community structures in large-scale networks. *Physical review E* **76**, 036106 (2007).

[110] Olsen, W. P. *The anti-corruption handbook: how to protect your business in the global marketplace* (John Wiley & Sons, 2010).

[111] Ramamoorti, S. The psychology and sociology of fraud: Integrating the behavioral sciences component into fraud and forensic accounting curricula. *Issues in Accounting Education* **23**, 521–533 (2008).

[112] Reichardt, J. & Bornholdt, S. Statistical mechanics of community detection. *Physical Review E* **74**, 016110 (2006).

[113] Rest, J. R., Narvaez, D., Thoma, S. J. & Bebeau, M. J. A neo-kohlbergian approach to morality research. *Journal of moral education* **29**, 381–395 (2000).

[114] Reuters. Wife Says She and Madoff Tried Suicide (2011). URL http://www.nytimes.com/2011/10/27/business/wife-says-she-and-madoff-tried-suicide.html.

[115] Romero, S. & Atlas, R. D. Worldcom's collapse: The Overview; Worldcom files for bankruptcy; Largest U.S. Case (2002). URL http://www.nytimes.com/2002/07/22/us/worldcom-s-collapse-the-overview-worldcom-files-for-bankruptcy-largest-us-case.html.

[116] Rosvall, M. & Bergstrom, C. T. An information-theoretic framework for resolving community structure in complex networks. *Proceedings of the National Academy of Sciences* **104**, 7327–7331 (2007).

[117] Savage, D., Zhang, X., Yu, X., Chou, P.& Wang, Q. Anomaly detection in online social networks. *Social Networks* **39**, 62–70 (2014).

[118] Schweitzer, M. E., Ordóñez, L. & Douma, B. Goal setting as a motivator of unethical behavior. *Academy of Management Journal* **47**, 422–432 (2004).

[119] Sezer, O., Gino, F. & Bazerman, M. H. Ethical blind spots: Explaining unintentional unethical behavior. *Current Opinion in Psychology* **6**, 77–81 (2015).

[120] Shalvi, S., Eldar, O. & Bereby-Meyer, Y. Honesty requires time (and lack of justifications). *Psychological science* **23**, 1264–1270 (2012).

[121] Shalvi, S., Gino, F., Barkan, R. & Ayal, S. Self-serving justifications doing wrong and feeling moral. *Current Directions in Psychological Science* **24**, 125–130 (2015).

[122] Shariff, A. F. & Norenzayan, A. Mean gods make good people: Different views of god predict cheating behavior. *The International Journal for the Psychology of Religion* **21**, 85–96 (2011).

[123] Shu, L. L., Gino, F. & Bazerman, M. H. Dishonest deed, clear conscience: When cheating leads to moral disengagement and motivated forgetting. *Personality and Social Psychology Bulletin* **37**, 330–349 (2011).

[124] Shu, L. L., Mazar, N., Gino, F., Ariely, D. & Bazerman, M. H. Signing at the beginning makes ethics salient and decreases dishonest self-reports in comparison to signing at the end. *Proceedings of the National Academy of Sciences* **109**, 15197–15200 (2012).

[125] Simon, H. A. A behavioral model of rational choice. *The quarterly journal of economics* 99–118 (1955).

[126] Slemrod, J. Cheating ourselves: The economics of tax evasion. *The journal of economic perspectives* **21**, 25–48 (2007).

[127] Smith, A. & Nicholson, J. S. *An Inquiry Into the Nature and Causes of the Wealth of Nations...* (T. Nelson and Sons, 1887).

[128] Smith, T. R. Low self-control, staged opportunity, and subsequent fraudulent behavior. *Criminal Justice and Behavior* **31**, 542–563 (2004).

[129] Srivastava, A., Kundu, A., Sural, S. & Majumdar, A. Credit card fraud detection using hidden markov model. *IEEE Transactions on dependable and secure computing* **5**, 37–48 (2008).

[130] Swiss Criminal Code, Art. 146. Schweizerisches Strafgesetzbuch (2014). URL `http://www.legislationline.org/documents/section/criminal-codes`.

[131] Tenbrunsel, A. E. & Messick, D. M. Ethical fading: The role of self-deception in unethical behavior. *Social Justice Research* **17**, 223–236 (2004).

[132] Thau, S., Derfler-Rozin, R., Pitesa, M., Mitchell, M. S. & Pillutla, M. M. Unethical for the sake of the group: Risk of social exclusion and pro-group unethical behavior. *Journal of Applied Psychology* **100**, 98 (2015).

[133] Thesaurus.com. Roget's 21st Century Thesaurus (2009). URL `http://www.thesaurus.com/browse/unethical`.

[134] Trevino, L. K. Ethical decision making in organizations: A person-situation interactionist model. *Academy of management Review* **11**, 601–617 (1986).

[135] Tversky, A. & Kahneman, D. Judgment under uncertainty: Heuristics and biases. In *Utility, probability, and human decision making*, 141–162 (Springer, 1975).

[136] University of Michigan Library. Academic Integrity in Social Work : Examples of Unethical Behavior (2016). URL `http://guides.lib.umich.edu/c.php?g=283365{\&}p=1887165`.

[137] Vohs, K. D., Mead, N. L. & Goode, M. R. The psychological consequences of money. *science* **314**, 1154–1156 (2006).

[138] Walker, L. J. & Pitts, R. C. Naturalistic conceptions of moral maturity. *Developmental psychology* **34**, 403 (1998).

[139] Wasserman, S. & Faust, K. *Social network analysis: Methods and applications*, vol. 8 (Cambridge university press, 1994).

[140] Wikipedia. Fraud (2016). URL https://en.wikipedia.org/wiki/Fraud.

[141] Wolfe, D. T. & Hermanson, D. R. The fraud diamond: Considering the four elements of fraud. *The CPA Journal* **74**, 38 (2004).

[142] Xu, Z. X. & Ma, H. K. How can a deontological decision lead to moral behavior? the moderating role of moral identity. *Journal of Business Ethics* 1–13 (2015).

[143] Your Dictionairy. Examples of Unethical Behavior (2016). URL http://examples.yourdictionary.com/examples-of-unethical-behavior.html.

[144] Zhang, T., Fletcher, P. O., Gino, F. & Bazerman, M. H. Reducing bounded ethicality. *Organizational Dynamics* **4**, 310–317 (2015).

4 A Comparative Analysis of Community Detection Algorithms on Artificial Networks[5]

Abstract

Many community detection algorithms have been developed to uncover the mesoscopic properties of complex networks. However how good an algorithm is, in terms of accuracy and computing time, remains still open. Testing algorithms on real-world network has certain restrictions which made their insights potentially biased: the networks are usually small, and the underlying communities are not defined objectively. In this study, we employ the Lancichinetti-Fortunato-Radicchi benchmark graph to test eight state-of-the-art algorithms. We quantify the accuracy using complementary measures and algorithms' computing time. Based on simple network properties and the aforementioned results, we provide guidelines that help to choose the most adequate community detection algorithm for a given network. Moreover, these rules allow uncovering limitations in the use of specific algorithms given macroscopic network properties. Our contribution is threefold: firstly, we provide actual techniques to determine which is the most suited algorithm in most circumstances based on observable properties of the network under consideration. Secondly, we use the mixing parameter as an easily measurable indicator of finding the ranges of reliability of the different algorithms. Finally, we study the dependency with network size focusing on both the algorithm's predicting power and the effective computing time.

4.1 Introduction

Relationships between constituents of complex systems (be it in nature, society, or technological applications) can be represented in terms of networks. In this portrayal, the elements

[5]**Author Statement:** This is a published work: Yang, Z., Algesheimer, R., & Tessone, C. J. (2016). A Comparative Analysis of Community Detection Algorithms on Artificial Networks. *Scientific Reports*, 6; doi: 10.1038/srep30750. Zhao Yang is the 1^{st} and corresponding author, René Algesheimer is the 2^{nd} author, and Claudio J. Tessone is the 3^{rd} author.

composing the system are described as nodes and their interactions as links. At the global level, the topology of these interactions – far from being trivial – is in itself of complex nature [1, 2]. Importantly, these networks further display some level of organisation at an intermediate scale. At this *mesoscopic* level, it is possible to identify groups of nodes that are heavily connected among themselves, but sparsely connected to the rest of the network. These interconnected groups are often characterised as *communities*, or in other contexts *modules*, and occur in a wide variety of networked systems [3, 4].

Detecting communities has grown into a fundamental, and highly relevant problem in network science with multiple applications. First, it allows to unveil the existence of a non-trivial internal network organisation at coarse grain level. This allows further to infer special relationships between the nodes that may not be easily accessible from direct empirical tests [5]. Second, it helps to better understand the properties of dynamic processes taking place in a network. As paradigmatic examples, spreading processes of epidemics and innovation are considerably affected by the community structure of the graph [6].

Taking into account its importance, it is not surprising that many community detection methods have been developed, using tools and techniques from variegated disciplines such as statistical physics, biology, applied mathematics, computer science, and sociology. All these methods aim at improving the identification of meaningful communities, while keeping as low as possible the computational complexity of the underlying algorithm. Clearly, these algorithms are based on slightly different definitions of community, and therefore the results are not always directly comparable. Further, in most real-world applications, a *ground truth* – i.e. a *unique* identification of nodes to communities – is simply non-existent, which makes it even more difficult to assess the reliability of the community detection procedures. To address these shortcomings and test the algorithms' reliability, different benchmarks have been developed.

Essentially, testing a community detection algorithm implies analysing computer-generated or real-world networks with a well defined community structure (a known ground truth) in order to obtain the community decomposition. One of the most used techniques is the GN benchmark (for Girvan & Newman [3]), which is a special case of the planted $l-$partition model [7] with

a prior specification of the number of nodes (128) and equally sized communities (4). When the expected number of links joining a node to others in different groups is smaller than 8, the four groups are strongly defined communities. In these conditions, a well functioning detection algorithm should be able to identify the communities in reasonable time. Different community detection algorithms can be compared based on their performances on the GN benchmark, which has already been done by Danon *et al.* [8]. However, there are several drawbacks to the GN benchmark: All nodes have the same expected degree, communities are separated in the same way, and the network is of an unrealistic small size.

It is a well established fact that most real complex networks are characterised by largely heterogeneous degree distributions [1, 2, 9] and heterogeneous community sizes [10, 11, 12]. For this reason, the GN benchmark cannot be considered as a good proxy for a real network. By consequence, in a newer stream of research [5, 13], the authors proposed an alternative benchmark, which is usually referred to as LFR (for Lancichinetti, Fortunato & Radicchi). This method introduces power-law distributions of degree and community size to the graphs to generalise the GN benchmark. The performances of most existing community detection algorithms are good on the GN benchmark. In contrast, the LFR benchmark presents a harder test for algorithms and makes it easier to unveil their limitations. It has been shown that the *mixing parameter*, which is defined as

$$\mu = \frac{\sum_i k_i^{ext}}{\sum_i k_i^{tot}} \tag{3}$$

is the most influential parameter in the LFR benchmark graphs [14]. Here k_i^{ext} and k_i^{tot} stand for the external degree of node i, i.e. the number of edges connecting it to others that belong to different communities, and the total degree of said node. Although it would be possible to define a mixing parameter for each node, it is assumed that μ is a global property and is the same for every node in the LFR benchmark. The reason here is to be consistent with the standard hypotheses of the planted l-partition model [15].

According to the definition of community in a strong sense, each node should have more

connections within the community than with the rest of the graph [16]. Therefore, for $\mu > 1/2$ communities in the strong sense disappear. However, it is worth to mention that Lancichinetti and Fortunato [15] found a weaker condition for community detection which can be applied to any version of the planted l-partition model: $\mu < (N - n_c^{max})/N$, where N is the total number of nodes, and n_c^{max} is the size of the largest community. In our study, although we stick to the strong definition of communities, we have also taken the general condition of μ into consideration (see Table 20).

In the following, we briefly review studies comparing community detection algorithms in chronological order [8, 5, 13, 15, 14, 17, 18] to highlight the research interests shift. In one of the early studies in comparing community detection algorithms, Danon *et al.* had tested ten algorithms on the GN benchmark7 [8] and collected estimates of how time complexity scales with network observables. However, the authors were not able to compare the actual computational effort as a result of the small sizes of graphs. Later on, Lancichinetti *et al.* had employed the LFR benchmark to measure the accuracy of two algorithms on undirected unweighted networks without overlapping communities [5] and two algorithms on directed weighted networks with overlapping communities [13]. Concurrently, the authors tested twelve different algorithms on the GN and LFR benchmarks, and random graphs. For the tests on the LFR benchmark, the authors had considered various parameters, including undirected unweighted graphs with non-overlapping communities, directed unweighted graphs with non-overlapping communities, undirected weighted graphs with non-overlapping communities, and undirected unweighted graphs with overlapping communities [15]. Orman and Labatut later tested five community detection algorithms on the LFR benchmark [14]. They measured the accuracy of algorithms and studied the properties of the LFR benchmark graphs. Later, Peel applied two algorithms on both weighted and unweighted networks with 100 nodes and examined the performance of algorithms developed for weighted networks against those for unweighted ones for different parts of the problem space [17]. Recently, Hric *et al.* compared the accuracy of eleven different algorithms on both the LFR benchmark and a collection of real world graphs with sizes vary from 34 to 5189809 nodes [18]. Overall, as an extension of the GN benchmark, the LFR has

drawn a lot of attention: Early, researchers employed small artificial and/or real world networks as benchmarks (e.g. the GN benchmark and the Zachary's karate club network) ; while nowadays people shifted towards the use of large stylised large artificial or real world networks with some kind of ground truth obtained from metadata information (e.g. the LFR benchmark and the DBLP collaboration network [19]). However, as of today, a detailed study of the dependency with the network size is missing as most of the existing studies include a few, selected, set of values of the number of nodes and the mixing parameter, and do not consider the real computing time needed to perform the analysis.

In this paper, we evaluate eight different state-of-the-art community detection algorithms available in the "igraph" package [20], which is a widely used collection of network analysis tools in R, Python, C and C++, on the LFR benchmark for undirected, unweighted graphs with non-overlapping communities. Details of the algorithms can be found in the methods section. Our contribution is threefold: First and foremost, we provide actual techniques to determine which is the most suited algorithm in most circumstances based on observable properties of the network under consideration. Secondly, we use the mixing parameter as an easily measurable indicator of finding the ranges of reliability of the different algorithms. Finally, we systematically study the dependency with network size focusing on both the algorithm's predicting power and the effective computing time.

4.2 Methods

In this section, we first describe in detail the procedure to obtain the benchmark networks used, then enumerate the community detection algorithms employed.

When comparing community detection algorithms, we can use either real or artificial network whose community structure is already known, which is usually termed as ground truth. Among the former, the celebrated Zachary's karate club [28] or the network of American college football teams [3] have been extensively used. Among the latter, the ones used more pervasively are the GN [3] and LFR [13] benchmarks. However, obtaining real networks to which a ground truth can be associated is not only difficult, but also costly in economic terms and time.

Due to the complexity of data collection and costs, real world benchmarks usually consist of small-sized networks. Further, since it is not possible to control all the different features of a real network (e.g. average degree, degree distribution, community sizes, etc.), the algorithms can only be tested – if resorting in this kind of graphs – on very specific cases with a limited set of features. In addition, the communities of real world networks are not always defined objectively or, in the best case, they rarely have a unique community decomposition. On the other hand, artificially generated networks can overcome most of these limitations. Given an arbitrary set of meso- or macroscopic properties, it is possible to generate randomly an ensemble of networks that respect them, in what is usually called generative models. However, as one of the most popular generative models, GN benchmark suffers from the fact that it does not show a realistic topology of the real network [29, 5] and it has very small network size. A recent strand of the literature on benchmark graphs tried to improve the quality of artificial networks by defining more realistic generative models: Lancichinetti *et al.* extended the GN benchmark by introducing power law degree and community size distributions [5]. Bagrow had employed the Barabási-Albert model [9] rather than the configuration model [30] to build up the benchmark graph [31]. Orman and Labatut proposed to use evolutionary preferential attachment model [32] for more realistic properties [33].

The first step to generate the LFR benchmark graph is to construct a network composed of N nodes, with average degree $\hat{k}$, maximum degree k_{max} and a power-law degree distribution with exponent α by using the configuration model. Once this step is finished, each node has a defined total degree. Then, given a power-law distribution of community sizes with exponent β, a set of community sizes is drawn (between arbitrarily chosen minimum and maximum values of community sizes that act as additional parameters). Nodes are then sequentially assigned to these communities. The mixing parameter μ, which represents the fraction of edges a node has with nodes belonging to other communities with respect to its total degree, is the most relevant value in terms of the community structure. To conclude the generative algorithm, edges are rewired in order to fit the mixing parameter, while preserving the degree sequence. This is achieved keeping fixed total degree of a node, the value of external degree is

modified so that the ratio of external degree over the total degree is close to the defined mixing parameter. The LFR model was initially proposed to generate undirected unweighted networks with mutually exclusive communities, and was extended to generate weighted and/or directed networks, with or without overlapping communities. In this study, we focus on the undirected unweighted networks with non-overlapping communities since most of the existing community detection algorithms are designed for this type of networks. The parameter values used in our computer-generated graphs are indicated in Table 20.

Table 20: Parameters of LFR benchmark graphs.

Parameter	Value
Number of nodes N	$233 \sim 31948$
Maximum degree	$0.1N$
Maximum community size	$0.1N$
Average degree	20
Degree distribution exponent	-2
Community size distribution exponent	-1
Mixing coefficient μ	[0.03, 0.75]

To deal with possible discrepancies in the network properties, we have randomly generated 100 network for every set of parameters. Due to the slow computing speed, Spinglass and Edge betweenness algorithms have been tested only on small networks with $N \leq 1000$.

In this paper, we have evaluated the most widely used, state-of-the-art community detection algorithms on the LFR benchmark graphs. In order to make the results comparable, and reproducible, we use the implementation of these algorithms shipped with the widely used "igraph" software package (Version 0.7.1) [20]. Here is the list of algorithms we have considered. For notation purposes when giving the computational complexity of the algorithms, the networks have N nodes and E edges.

Edge betweenness

This algorithm was introduced by Girvan & Newman [3]. To find which edges in a network exist most frequently between other pairs of nodes, the authors generalised Freeman's betweenness

centrality [34] to edges betweenness. The edges connecting communities are then expected to have high edge betweenness. The underlying community structure of the network will be much clear after removing edges with high edge betweenness. For the removal of each edge, the calculation of edge betweenness is $\mathcal{O}(EN)$; therefore, this algorithm's time complexity is $\mathcal{O}(E^2N)$ [3].

Fastgreedy

This algorithm was proposed by Clauset *et al.* [12]. It is a greedy community analysis algorithm that optimises the modularity score. This method starts with a totally non-clustered initial assignment, where each node forms a singleton community, and then computes the expected improvement of modularity for each pair of communities, chooses a community pair that gives the maximum improvement of modularity and merges them into a new community. The above procedure is repeated until no community pairs merge leads to an increase in modularity. For sparse, hierarchical, networks the algorithm runs in $\mathcal{O}(N \log^2(N))$ [12].

Infomap

This algorithm was proposed by Rosvall *et al.* [35, 36]. It figures out communities by employing random walks to analyse the information flow through a network [17]. This algorithm starts with encoding the network into modules in a way that maximises the amount of information about the original network. Then it sends the signal to a decoder through a channel with limited capacity. The decoder tries to decode the message and to construct a set of possible candidates for the original graph. The smaller the number of candidates, the more information about the original network has been transferred. This algorithm runs in $\mathcal{O}(E)$ [37].

Label propagation

This algorithm was introduced by Raghavan *et al.* [38]. It assumes that each node in the network is assigned to the same community as the majority of its neighbours. This algorithm starts with initialising a distinct label (community) for each node in the network. Then, the nodes in the

network are listed in a random sequential order. Afterwards, through the sequence, each node takes the label of the majority of its neighbours. The above step will stop once each node has the same label as the majority of its neighbours. The computational complexity of label propagation algorithm is $\mathcal{O}(E)$ [38].

Leading eigenvector

This algorithm was proposed by Newman [39]. The heart of this algorithm is the spectral optimisation of modularity by using the eigenvalues and eigenvectors of the modularity matrix. First, the leading eigenvector of the modularity matrix is calculated, and then the graph is split into two parts in a way that modularity improvement is maximised based on the leading eigenvector. After that, the modularity contribution is calculated at each step in the subdivision of a network. It stops once the value of the modularity contribution is not positive. Its computational complexity of each graph bipartition is $\mathcal{O}(N(E+N))$, or $\mathcal{O}(N^2)$ on a sparse graph [40].

Multilevel

This algorithm was introduced by Blondel *et al.* [25]. It is a different greedy approach for optimising the modularity with respect to the Fastgreedy method. This method first assigns a different community to each node of the network, then a node is moved to the community of one of its neighbours with which it achieves the highest positive contribution to modularity. The above step is repeated for all nodes until no further improvement can be achieved. Then each community is considered as a single node on its own and the second step is repeated until there is only a single node left or when the modularity can't be increased in a single step. The computational complexity of the Multilevel algorithm is $\mathcal{O}(N\log N)$ [40].

Spinglass

This algorithm was first proposed by Reichardt & Bornholdt [41]. It is based on the Potts model [42]. The basic principle of the method is that edges should connect nodes of the same spin state (community, in the current context), whereas nodes of different states (belonging to different

communities) should be disconnected. Therefore, the aim of this algorithm is to find the ground state of a spin glass model with a Potts Hamiltonian. Simulated annealing [43] has been used to minimise the system's free energy [44]. In a sparse graph, the computational complexity of this algorithm is approximately $\mathcal{O}(N^{3.2})$ [45].

Walktrap

This algorithm was proposed by Pon & Latapy [46]. It is a hierarchical clustering algorithm. The basic idea of this method is that short distance random walks tend to stay in the same community. Starting from a totally non-clustered partition, the distances between all adjacent nodes are computed. Then, two adjacent communities are chosen, they are merged into a new one and the distances between communities are updated. This step is repeated $(N-1)$ times, thus the computational complexity of this algorithm is $\mathcal{O}(EN^2)$. For sparse networks the computational complexity is $\mathcal{O}(N^2\log(N))$ [40].

We have employed virtual machines to implement all the computation. For each network size and for each algorithm, a virtual machine is created using a pre-defined installation that guarantees the same execution environment conditions. The installation is tuned to guarantee that each virtual machine makes use of an entire physical node, and, at the same time, that all physical nodes where the virtual machines will be hosted have the very same hardware specifications. The workload distribution and collection for the results are commanded by a master-slave approach.

4.3 Results

In this section, we compare the results of community detection algorithms in terms of accuracy and computing time. The former is defined as a measure of similarity between the modular structure generated by the LFR benchmark $\mathscr{P}$ (see Methods Section) and the partition identified by the respective community detection algorithms $\bar{\mathscr{P}}$. The latter is the real computing time needed to perform the community detection. This section is organised as follows: First, by

employing the LFR generative model, we unveil the relationship between the mixing parameter and the accuracy of the community detection algorithms. Accuracy is measured in two different, complementary ways: The normalised mutual information [8], and the ratio between the number of detected communities and the number of communities given by the LFR generating model. Then, we measure the computing time of community detection algorithms and show the relationship between the mixing parameter and the computing time. We then present the mixing parameter as computed from the communities detected by the different algorithms as a function of the input mixing parameter. Last, we present the comparisons of community detection algorithms in terms of accuracy and computing time as a function of network sizes.

4.3.1 The role of the network mixing parameter on accuracy and computing time

First, we study the accuracy of the community detection algorithms as a function of the mixing parameter μ. To measure the accuracy we have employed the *normalised mutual information*, i.e., NMI. This is a measure borrowed from information theory which has been regularly used in papers comparing community detection algorithms [13].

Defining a *confusion matrix* **N**, where the rows correspond to the 'real' communities, and the columns correspond to the 'found' communities. The element of **N**, N_{ij}, is the number of nodes in the real community i that appear in the j-th detected community. The *normalised mutual information* is then [8]

$$I(\mathscr{P},\bar{\mathscr{P}}) = \frac{-2\sum_{i=1}^{C}\sum_{j=1}^{\bar{C}} N_{ij}\log(N_{ij}N/N_{i\circ}N_{\circ j})}{\sum_{i=1}^{C} N_{i\circ}\log(N_{i\circ}/N)+\sum_{j=1}^{\bar{C}} N_{\circ j}\log(N_{\circ j}/N)} \tag{4}$$

where the number of communities given by the LFR model is denoted by C and the number of communities detected by the algorithm is denoted by $\bar{C}$. The sum over the i-th row of **N** is denoted $N_{i\circ}$ and the sum over the j-th column is denoted $N_{\circ j}$. If the estimated communities are identical to the real ones, $I(\mathscr{P},\bar{\mathscr{P}})$ equals to 1. If the partition found by the algorithm is totally independent from the real partition, $I(\mathscr{P},\bar{\mathscr{P}})$ vanishes.

As pointed out in Ref. [21], the mutual information can be normalised in different ways. These different normalisation methods are sensitive to different partition properties and have different theoretical properties [21, 22, 23]. To get a better overview of the accuracy, we have calculated the NMI by using all these five different definitions (cf. SI). We conclude that in the current study different normalisation procedures provide qualitatively similar behaviours. Just for the sake of brevity, and consistently with Danon *et al.* [8], we report in this section only I_{sum} (i.e. normalisation by the arithmetic mean). The results of the other NMIs are shown in "Supplementary Information of Chapter 4".

The results are shown in Figure 10. Each panel presents the accuracy of a given community detection algorithm and is subdivided into two plots: The lower axis depict the average value of NMI and the upper ones contain the standard deviation of the measures when repeated over 100 different network realisations. Most of the algorithms can uncover well the communities when the mixing parameter μ is small, as it is apparent from the large values of I in the limit $\mu \to 0$. The accuracy of algorithms decreases, then, with increasing values of both network size and μ. Different algorithms behave differently: the accuracy of Fastgreedy algorithm decreases monotonically, in a smooth fashion and has a very small standard deviation along all the range (Panel (a), Figure 10). Whereas that of Leading eigenvector algorithm falls rapidly even with small value of μ (Panel (c), Figure 10). All the other algorithms display abrupt changes of behaviour: their performances remain relatively stable before a turning point where the NMI drops very fast as a function of μ. The changes of behaviour are usually around $\mu = 1/2$, which corresponds to the strong definition of community [16]. Interestingly, Label propagation and Edge betweenness algorithms have turning points smaller than said value; while Infomap, Multilevel, Walktrap, and Spinglass algorithms have turning points greater than $\mu = 1/2$. We have also noticed that for the Infomap algorithm the normalised mutual information has a point of discontinuous behaviour at around $\mu \cong 0.55$. On the other hand, for Label propagation, I vanishes around $\mu \cong 0.5$ falling in a continuous fashion. This supports the conjecture that Infomap displays a first order phase transition as a function of the mixing parameter, while Label propagation algorithm may have a second order one. Nonetheless, we have not performed

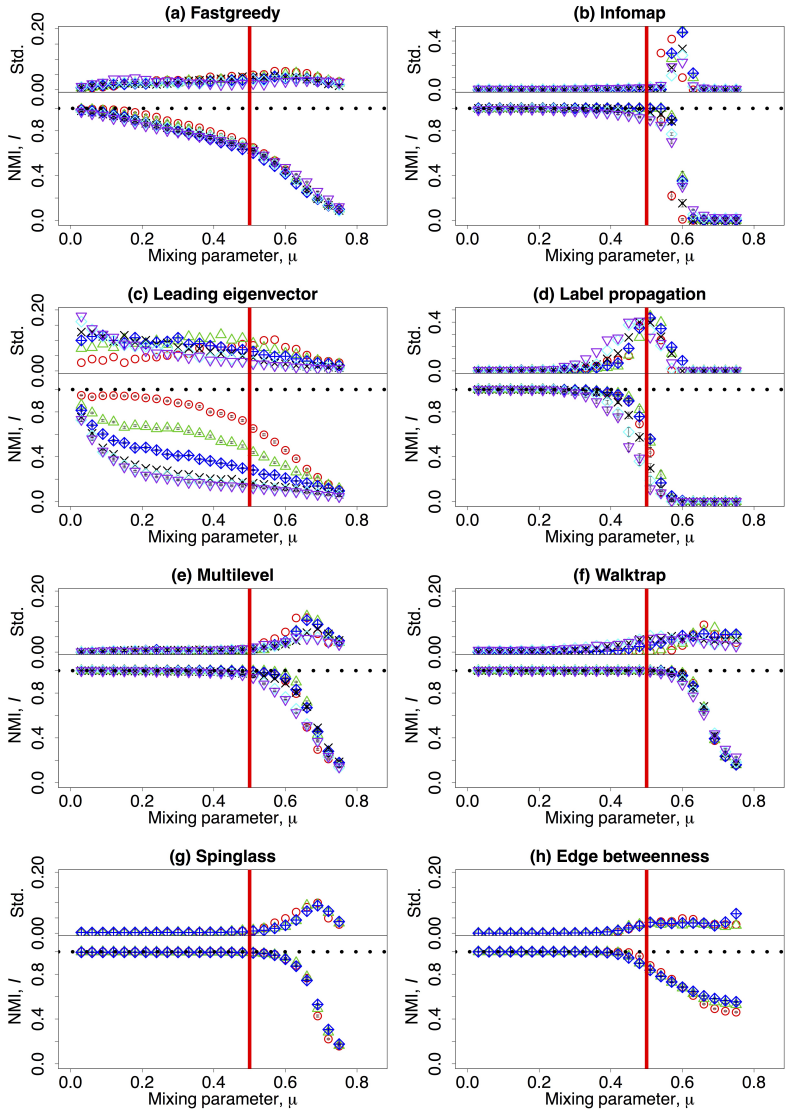
an exhaustive analysis on the matter to systematically analyse the existence (or not) of critical points. Further studies concerning the properties of these points are definitely needed.

Network size also plays the role here that a larger network size will lead to loss of accuracy at a lower value of μ. For small enough networks ($N \leq 1000$), Infomap, Multilevel, Walktrap, and Spinglass outperform the other algorithms with higher values of I and very small standard deviations, which shows the repeatability of the partitions detected. Besides, the turning point for accuracy is after $\mu = 1/2$. For larger networks ($N > 1000$), Infomap, Multilevel and Walktrap algorithms have relatively better accuracies and smaller standard deviations. Label propagation algorithm has much larger standard deviations such that its outputs are not stable. Due to the long computing time, Spinglass and Edge betweenness algorithms are too slow to be applied on large networks.

Second, we study how well the community detection algorithms reproduce the number of communities. To do so, we compute the ratio $\bar{C}/C$ as a function of the mixing parameter. $\bar{C}$ is the average number of detected communities delivered by the different algorithms when repeated over 100 different network realisations. C is the average real number of communities provided by the LFR benchmark on the same 100 networks. If $\bar{C}/C = 1$, the community detection algorithms are able to estimate correctly the number of communities. It is important to remark that this parameter has to be analysed together with the normalised mutual information because the distribution of community sizes is very heterogeneous. With respect to the networks generated by the LFR model, for small network sizes the real number of communities is stable for all values of μ, while for larger network sizes ($N > 1000$), C grows up to $\mu \gtrapprox 0.2$ and then it saturates.

The results for the ratio $\bar{C}/C$ as a function of the mixing parameter are shown in Figure 11 on a *log-linear* scale for all the panels. The Fastgreedy algorithm constantly underestimates the number of communities, and the results worsen with increasing network size and μ (Panel (a), Figure 11). For $\mu \lessapprox 0.55$, the Infomap algorithm delivers the correct number of communities of small networks ($N \lessapprox 1000$), and overestimates it for larger ones. For $\mu \gtrapprox 0.55$, this algorithm fails to detect any community at all for small networks and all nodes are partitioned into a single

Figure 10: (lower row) The mean value of normalised mutual information depending on the mixing parameter μ. (upper row) The standard deviation of the NMI as a function of μ.

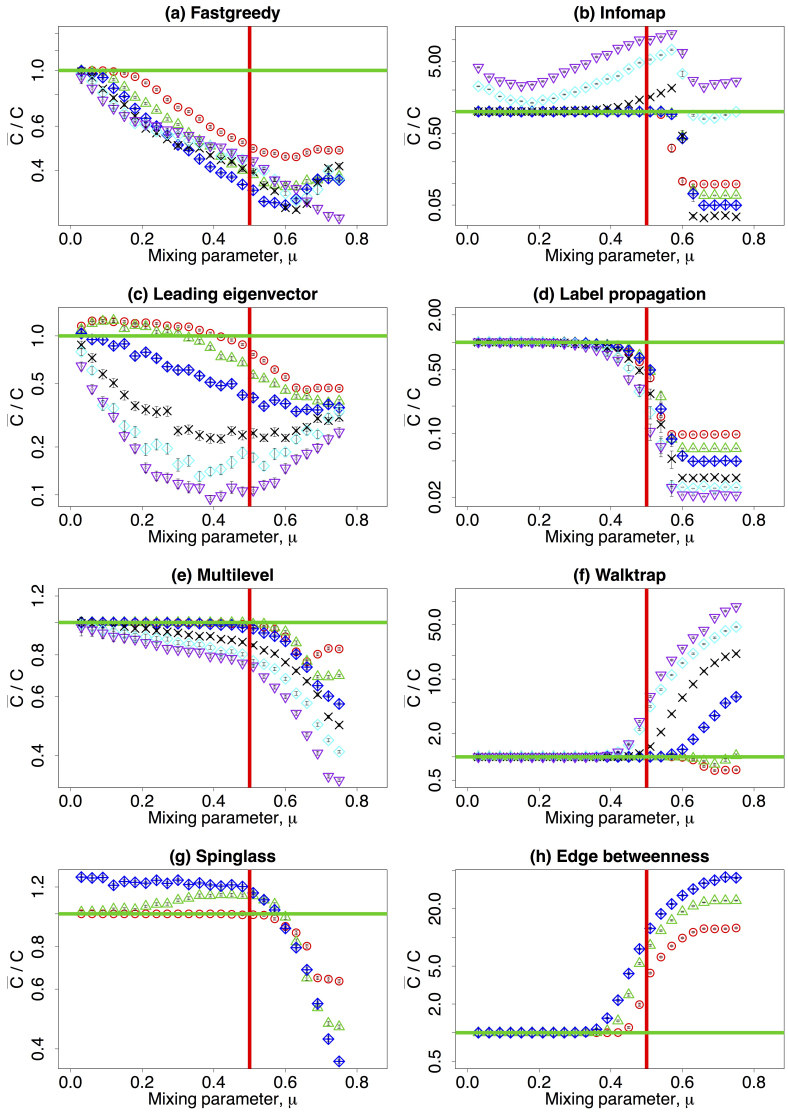


Different colours refer to different number of nodes: red ($N = 233$), green ($N = 482$), blue ($N = 1000$), black ($N = 3583$), cyan ($N = 8916$), and purple ($N = 22186$). Please notice that the vertical axis on the subfigures might have different scale ranges. The vertical red line corresponds to the strong definition of community, i.e. $\mu = 0.5$. The horizontal black dotted line corresponds to the theoretical maximum, $I = 1$. The other parameters are described in Table 20.

community (Panel (b), Figure 11). The leading eigenvector algorithm slightly overestimates the number of communities of small networks and the prediction worsens with increasing μ. Moreover, it underestimates the number of communities in large networks and even the behaviour do not change monotonically with μ (Panel (c), Figure 11). The Label propagation algorithm is able to deliver the correct number of communities with small values of μ regardless of the network size. However, in the range $0.3 \lessapprox \mu \lessapprox 0.6$, it underestimates the number of communities and the prediction worsens with increasing network size and μ. For $\mu \gtrapprox 0.6$, this algorithm fails to detect any community and all nodes are placed into the same community (Panel (d), Figure 11). It is apparent that the Mutilevel algorithm constantly underestimates the number of communities and such behaviour worsens with increasing network size and μ (Panel (e), Figure 11). In Figure 11, Panel (f), for $\mu \lessapprox 0.4$, the Walktrap algorithm delivers the correct number of communities regardless of network sizes, although the change of behaviour at which the prediction is correct depends on system size. For $\mu \gtrapprox 0.4$, this algorithm behaves differently depending on network size: it slightly underestimates the number of communities of small networks and significantly overestimates it for large ones. For $\mu \lessapprox 0.6$, the Spinglass algorithm constantly overestimates the number of communities, and its prediction worsens with network size. When $\mu \gtrapprox 0.6$, it fails and tends to put nodes into a few giant communities (Panel (g), Figure 11). The Edge betweenness algorithm is able to deliver the correct number of communities for $\mu \lessapprox 0.4$ regardless of network size. It overestimates C for $\mu \gtrapprox 0.4$ and the accuracy of the prediction worsens with increasing network size (Panel (h), Figure 11). Overall, for $\mu \lessapprox 1/2$, Infomap, Leading eigenvector, Multilevel, Spinglass, and Edge betweenness algorithms are able to deliver a reasonable estimator of the number of communities for small networks, while the number of communities obtained by Label propagation and Walktrap algorithms are relatively close to the real value regardless of network size. For $\mu \gtrapprox 1/2$, all the algorithms are much worse at detecting the correct number of communities, and among all the algorithms, Multilevel, Walktrap, and Spinglass algorithms have better outputs when the network sizes are small.

Third, we turn to the real computing time of the algorithms. This measure is usually represented in theoretical estimations as a function of the number of nodes and edges. However, the

Figure 11: The mean value of the estimated number of communities delivered by different algorithms over the real number of communities given by the LFR benchmark, i.e., $\bar{C}/C$, dependent on the mixing parameter μ on a *log-linear* scale.



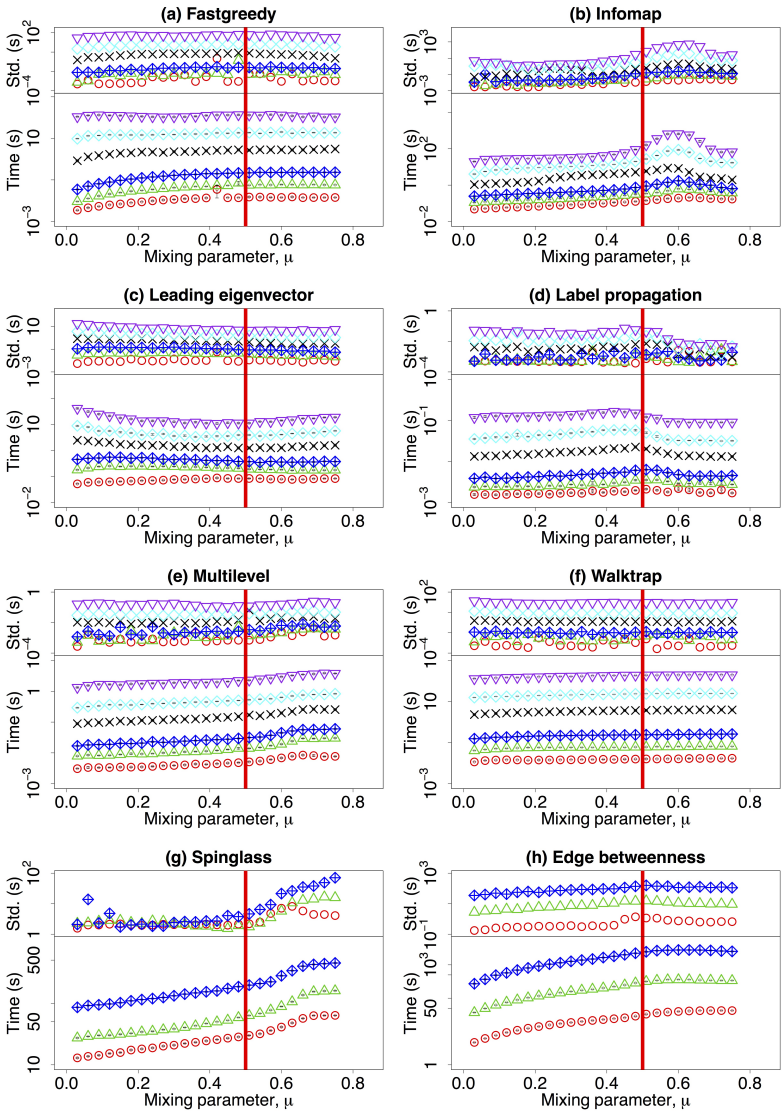
Different colours refer to different number of nodes: red ($N = 233$), green ($N = 482$), blue ($N = 1000$), black ($N = 3583$), cyan ($N = 8916$), and purple ($N = 22186$). Please notice that the vertical axis might have different scale ranges. The vertical red line corresponds to the strong definition of community where $\mu = 0.5$ and the horizontal green line represents the case that $\bar{C} = C$. The other parameters are described in Table 20.

real computing time may be also affected by the structure of the network. Given the number of nodes and a fixed average degree, we illustrate the computing time as a function of the mixing parameter. The results are shown in Figure 12 on *log-linear* scale. Each panel presents the computing time of a given community detection algorithm and it is subdivided in two plots: the lower one depicts the average computing time, while the upper sub-panel contains the standard deviation of the computing time when repeated over 100 different network realisations. Some algorithms barely depend on the mixing parameter. This is not the case for Multilevel, Spinglass, and Edge betweenness algorithms (Panel (e), (g), and (h), Figure 12). There is a slight dependency for Infomap algorithm that cannot be disregarded (Panel (b), Figure 12). The decrease of computing time for Infomap, Leading eigenvector, and Label propagation algorithms (Panel (b), (c), and (d), Figure 12) are accompanied with the significant worsening of NMI and $\bar{C}/C$ in Figures 10 and 11. Among all the algorithms, Label propagation and Multilevel algorithms are much faster than the others (Panel (d), and (e), Figure 12), while Spinglass and Edge betweenness are the slowest ones (Panel (g) and (h), Figure 12).

4.3.2 The observed mixing parameter

Unlike the number of nodes in a network, the exact value of the mixing parameter of a graph is unobservable if ground truth is unavailable for the community assignment of nodes. In this section, we study the mixing parameter delivered by the community detection algorithms $\bar{\mu}$ as a function of the mixing parameter μ (see Eq. 3). The results of the different algorithms are shown in the different panels of Figure 13. Each panel is subdivided in two plots: the lower has the average computed value of $\bar{\mu}$, while the upper sub-panel contains the standard deviation of the measures when repeated over 100 different network realisations. All algorithms have a linear (identity) relationship between $\bar{\mu}$ and μ except for the Leading eigenvector algorithm, which overshoots the results (Panel (c), Figure 13). Most of the algorithms display a turning point where the estimation of $\bar{\mu}$ breaks down. For the Fastgreedy, Multilevel, Walktrap, Spinglass, and Edge betweenness algorithms, $\bar{\mu}$ changes in a smooth fashion (Panel (a), (e), (f), (g), and (h), Figure 13). For the Infomap and Label propagation algorithms, the estimated mixing

Figure 12: (lower row) The mean value of the computing time of the community detection algorithms (in seconds) dependent on the mixing parameter μ on a *log-linear* scale. (upper row) The standard deviation of the measures on a *log-linear* scale.



Different colours refer to different number of nodes: red ($N = 233$), green ($N = 482$), blue ($N = 1000$), black ($N = 3583$), cyan ($N = 8916$), and purple ($N = 22186$). Please notice that the vertical axis might have different scale ranges. The vertical red line corresponds to the strong definition of community where $\mu = 0.5$. The other parameters are described in Table 20.

parameter $\bar{\mu}$ has a steep change at around $\mu \cong 0.55$ and $\mu \cong 0.5$, separately (Panel (b), and (d), Figure 13).

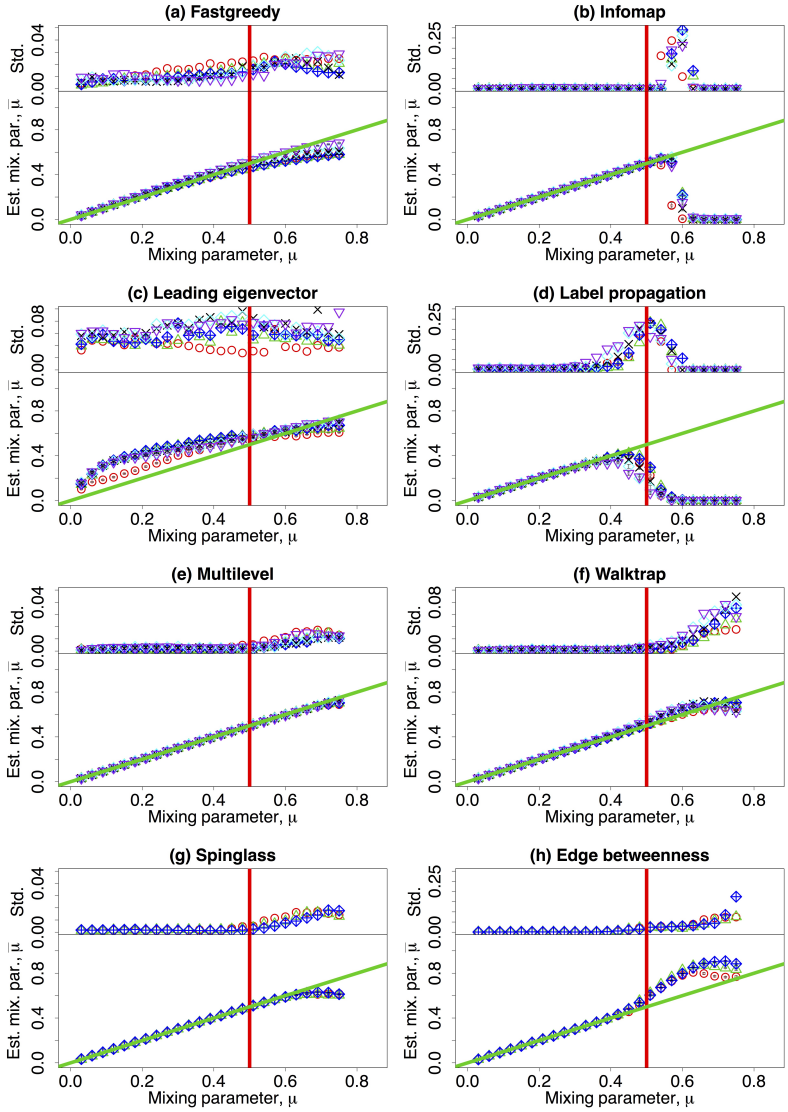
Overall, the mixing parameter obtained by the algorithms $\bar{\mu}$ fits well with the real mixing parameter at small value of μ, but it differs from the real value with increasing μ. For certain algorithms, the estimation fails completely for larger values of μ (Infomap, Label propagation), and for the others it is either overestimated (Edge betweenness) or slightly underestimated (Fastgreedy, Walktrap, Spinglass). Remarkably, in the Multilevel algorithm, the estimation is very accurate for values as large as $\mu = 0.75$ for all network sizes analysed.

4.3.3 The role of network size

So far we have only discussed the role of the mixing parameter μ to the accuracy and the computing time of community detection algorithms. Now, as an important ingredient, we consider the effect of network size. In our definition of the benchmark graphs, with a fixed average degree, network size can be represented as the number of nodes in the network. The results are shown in Figure 14 on a *linear-log* scale. Each of them presents the accuracy of a given community detection algorithms and is subdivided in two plots: one for the computed value of NMI and the upped sub-panel contains the standard deviation of the measures when repeated over 100 different network realisations. Most of the algorithms can well uncover the communities when $\mu \lessapprox 0.2$. In this case, the detecting abilities of Fastgreedy, Infomap, Label propagation, Multilevel, Walktrap, Spinglass and Edge betweenness algorithms are independent of network size (Panel (a,b,d-h), Figure 14). For Leading eigenvector, the accuracies decrease smoothly with network size (Panel (c), Figure 14). For very large $\mu \gtrapprox 0.75$, most of the algorithms fail to detect the community structure except for the Walktrap and Edge betweenness algorithms and the accuracy barely depends on network size. In the intermediate region of μ, NMI is usually decreasing with network size and μ.

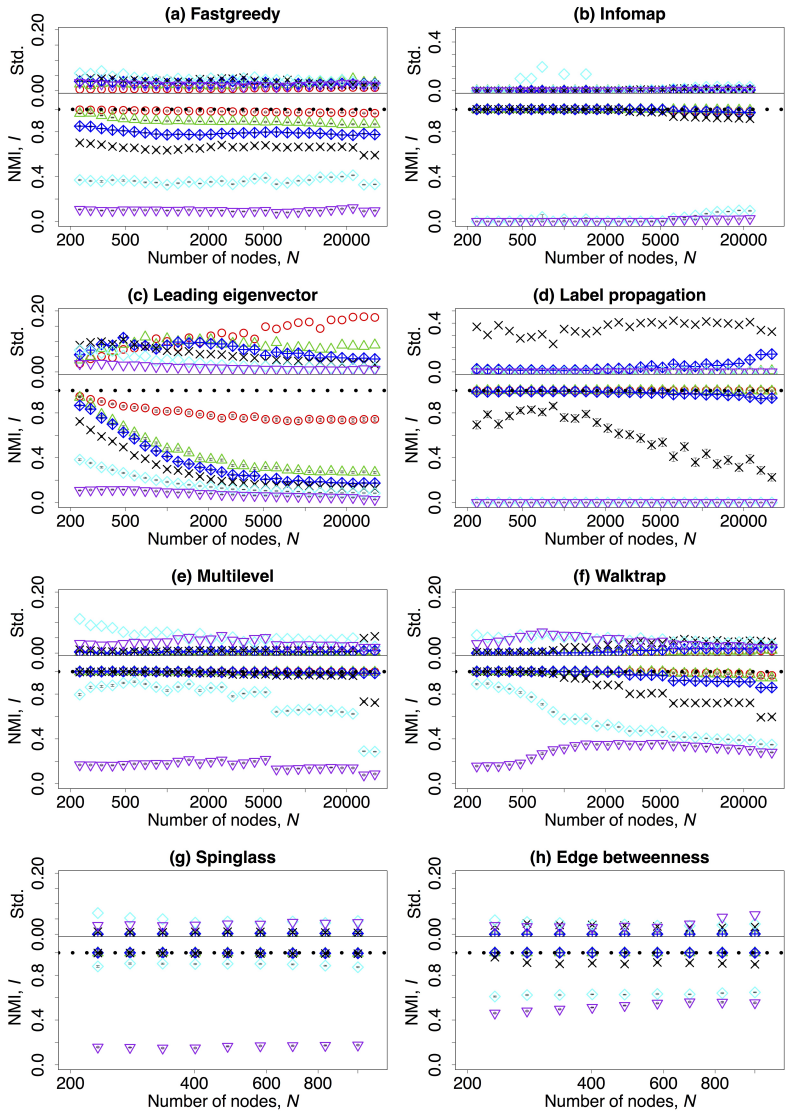
Finally, we present the computing time as a function of the network size. The results are represented in Figure 15 on a *log-log* scale. Each panel presents the computing time of a given community detection algorithms and is subdivided in two plots: one for the measured value

Figure 13: (lower row) The mean value of the mixing parameter estimated by the community detection algorithms $\bar{\mu}$ dependent on the mixing parameter μ. (upper row) The standard deviation of $\bar{\mu}$ dependent on μ.



Different colours refer to different number of nodes: red ($N = 233$), green ($N = 482$), blue ($N = 1000$), black ($N = 3583$), cyan ($N = 8916$), and purple ($N = 22186$). Please notice that the vertical axis on the subfigures might have different scale ranges. The vertical red line corresponds to the strong definition of community where $\mu = 0.5$. The green line $y = x$ corresponds to the case which $\bar{\mu} = \mu$. The other parameters are described in Table 20.

Figure 14: (lower row) The mean value of normalised mutual information dependent on the number of nodes N in the benchmark graphs on a *linear-log* scale. (upper row) The standard deviation of the normalised mutual information dependent on N on a *linear-log* scale.



Different colours refer to different values of the mixing parameter: red ($\mu = 0.03$), green ($\mu = 0.18$), blue ($\mu = 0.33$), black ($\mu = 0.48$), cyan ($\mu = 0.63$), and purple ($\mu = 0.75$). Please notice that the vertical axis on the subfigures might have different scale ranges. The horizontal black dotted line corresponds to $I = 1$. Due to the computing speed, Spinglass and Edge betweenness algorithms have been tested only on networks with $N \leq 1000$, and Infomap algorithm has been tested on networks with $N \leq 22186$. The other parameters are described in Table 20.

Table 21: Indexes of the exponential function $T \propto N^{\alpha}$ with the corresponding adjusted R-squared values.

	Fastgreedy	Infomap	Leading eigenvector	Label propagation
α	2.048 [0.006]	1.421 [0.009]	1.123 [0.005]	0.959 [0.005]
R^2	0.956	0.933	0.951	0.947
	Multilevel	Walktrap	Spinglass	Edge betweenness
α	1.126 [0.003]	2.04 [0.002]	1.282 [0.013] ()	2.915 [0.005]
R^2	0.957	0.962	0.867	0.884

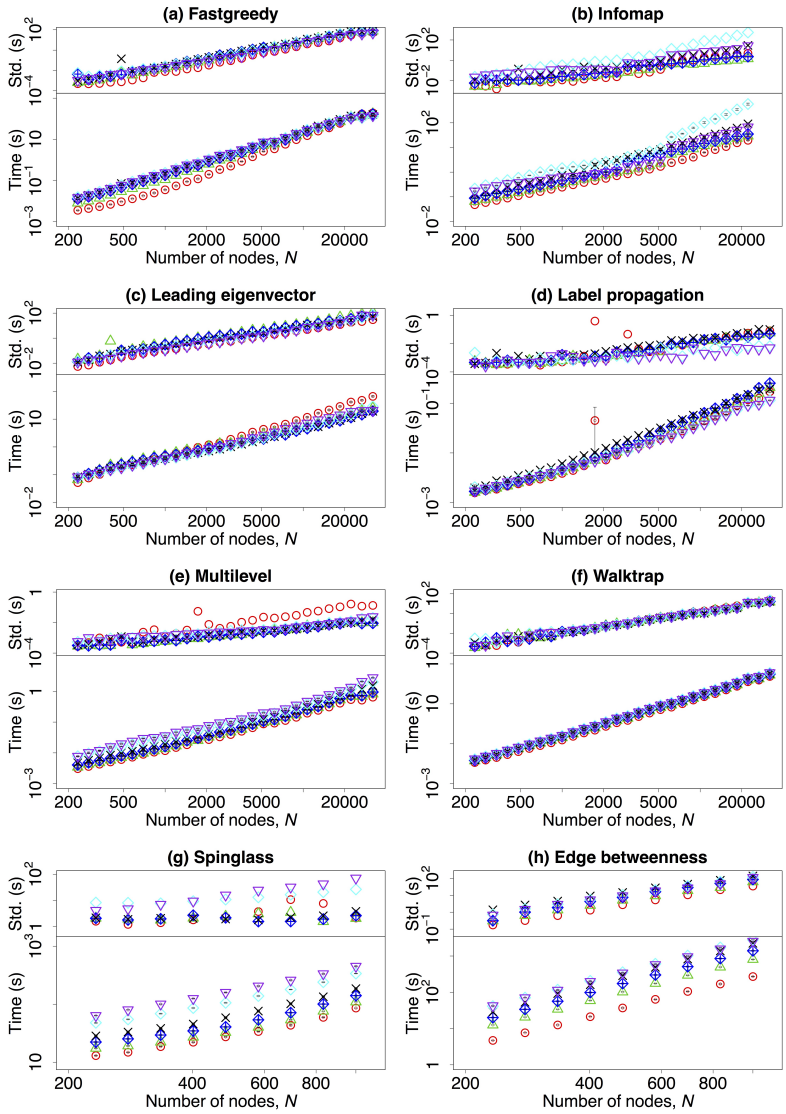
The standard errors are listed in brackets. All the results are statistically significant at the significance level of 0.05. Spinglass and Edge betweenness algorithms have been tested only on small networks with $N \leq 1000$, there might be some biases in the indexes of these two methods.

of computing time in second and the upped sub-panel contains the standard deviation of the measures when repeated over different network realisations. In the *log-log* scale, there is a significant linear correlation between the computing time and the network size. To further compare the computing speed of every algorithm, we have fitted the curves according to the exponential function $T \propto N^{\alpha}$. The fitted α together with the corresponding adjusted R-squared values are listed in Table 21. Only algorithms with small α can be applied to large networks. Overall, Label propagation algorithm is the method that scales best on network size; at the same time, Leading eigenvector, and Multilevel algorithms also have reasonable computation speeds on large networks. Fastgreedy, Infomap, Walktrap, and Spinglass algorithms scale much worse than the previous ones, and Edge betweenness algorithm is only suitable for small networks (with an almost cubic relation between network size and computing time).

4.4 Discussion

Traditionally, the aim of community detection in graphs has been to identify the modules by only using the information encoded in the graph topology [4]. In this study we have performed a comparative analysis of the accuracy and computing time of eight different community detection algorithms available in the "igraph" package. Each algorithm has been tested on a set of

Figure 15: (lower row) The mean value of the computing time of the community detection algorithms (in seconds) dependent on the number of nodes in the benchmark graphs on a *log-log* scale. (upper row) The standard deviation of the computing time on a *log-log* scale.



Different colours refer to different values of the mixing parameter: red ($\mu = 0.03$), green ($\mu = 0.18$), blue ($\mu = 0.33$), black ($\mu = 0.48$), cyan ($\mu = 0.63$), and purple ($\mu = 0.75$). Please notice that the vertical axis might have different scale ranges. Due to the computing speed, Spinglass and Edge betweenness algorithms have been tested only on networks with $N \leq 1000$, and Infomap algorithm has been tested on networks with $N \leq 22186$. The other parameters are described in Table 20.

LFR benchmark graphs [5, 13]. The size of the benchmark graphs varies from approximately 200 to 32,000 nodes. With a fixed average degree, we have changed the structure of networks by using different values of the mixing parameter μ.

In this study, the limited network sizes considered here pose no challenge for modern day computers in terms of Random-Access Memory (RAM). Therefore, the memory consumption is not analysed here. However, it is worth mentioning that the maximal memory consumption could be crucial for larger scale networks: if one algorithm is implemented in a way that it needs more memory for the optimal calculation, then it can easily happen that the process slows down for large networks due to low available RAM, or it switches to a suboptimal implementation, which needs less memory. A previous study showed [24] that (theoretically) many community detection methods have minimum memory consumption needs that scale linearly with the size of the graph $\mathcal{O}(2m+2n)$, where m is the number of edges and n is the number of nodes. In practice, many of them need at least $\mathcal{O}(2m+3n)$ in case of unweighted undirected graphs and when the Yale sparse matrix format is used [24].

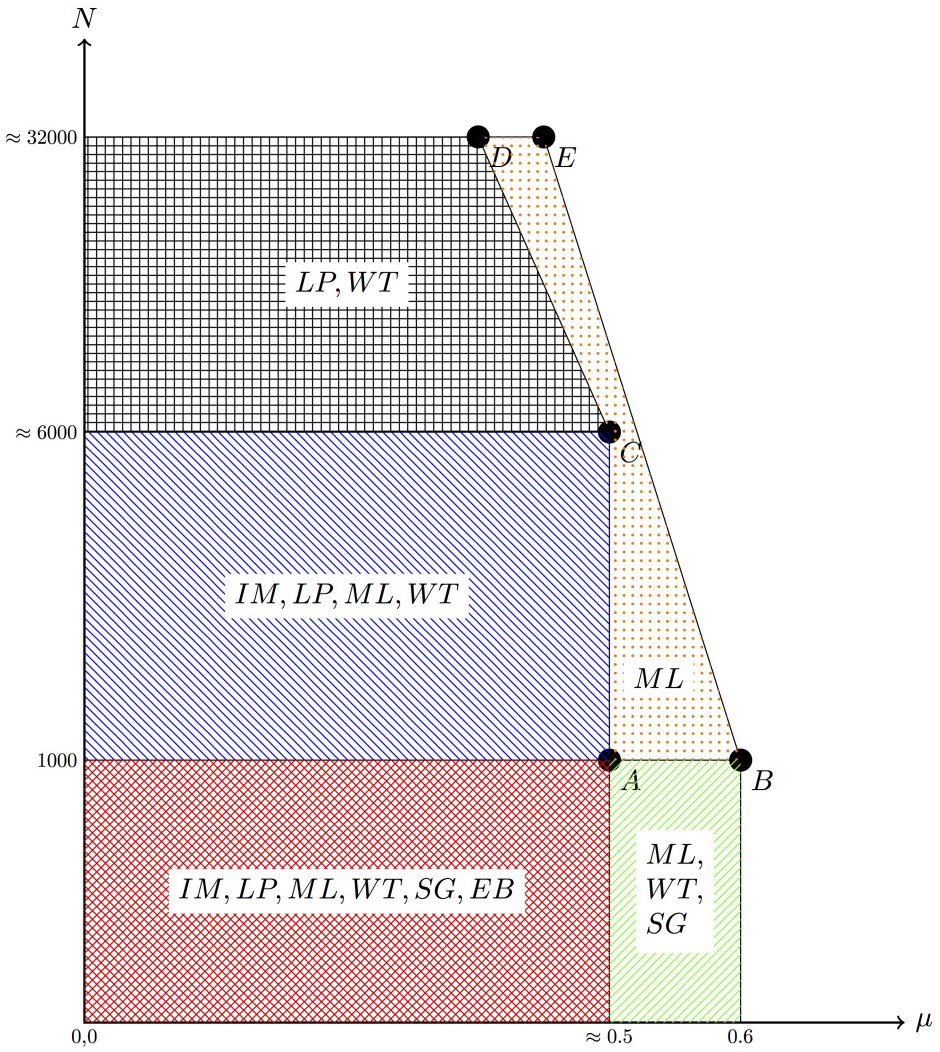
Our results indicate that by taking both accuracy and computing time into account, the Multilevel algorithm, which was proposed by Blondel *et al.* [25], outperforms all the other algorithms on the set of benchmarks we have examined (although the modularity-based methods are known to suffer from the resolution limit of modularity [26]). We can further apply the results in three aspects: First, since the computing time is not relevant for small networks, one should choose algorithms based on their accuracies. Among all the algorithms, Infomap, Label propagation, Multilevel, Walktrap, Spinglass, and Edge betweenness algorithms are able to successfully uncover the structure of small networks when the mixing parameter μ is small. With increasing value of μ, Infomap, Label propagation, and Edge betweenness algorithms' accuracies drop for smaller values of μ than Multilevel, Walktrap, and Spinglass algorithms. Second, for large networks, one should first choose algorithms which are able to detect the organisation of nodes in a reasonable time. In this sense, Infomap, Label propagation, Multilevel, and Walktrap algorithms are the *a priori* choices. After that, by taking the accuracy into account, Multilevel is superior to the other algorithms as it displays a performance drop for a

larger value of the mixing parameter μ. Importantly, the exact value of the mixing parameter of a graph is usually unobservable. To get a rough idea about the value of μ, one may employ either the Spinglass or the Multilevel algorithm. Limited by the computing time required, Spinglass algorithm cannot be applied on large networks.

Based on the previous results, and taking into account both factors, accuracy and computing time, it is possible to suggest under which situations to use each algorithm depending sorely on topological properties of the network under study. Our recommendations for the use of community detection algorithms are summarised in Figure 16. In the first region, $\mu \lessapprox 0.5$ and the network size is small, $N \lessapprox 1000$. There, most of the communities detection algorithms tested give accurate results (and the computing time is affordable): Infomap, Label propagation, Multilevel, Walktrap, Spinglass, and Edge betweenness can all be used in a trustworthy fashion. A second region has a relatively larger value of μ ($0.5 \lessapprox \mu \lessapprox 0.6$), and equally small sizes of network $N \lessapprox 1000$. There, it is possible to use Multilevel, Walktrap, and Spinglass algorithms. A third region encompasses again smaller values of mixing parameter ($\mu \lessapprox 0.5$) but an intermediate number of nodes ($1000 \lessapprox N \lessapprox 6000$). In this region, the best choices are Infomap, label propagation, Multilevel, and Walktrap algorithms. With increasing number of nodes in the networks ($6000 \lessapprox N \lessapprox 32000$), Infomap and Multilevel algorithm are very likely to provide the wrong number of communities and therefore they are no longer suitable in the fourth region. The last region has the highest requirement for the community detection algorithms. None of the algorithms performs very well in this region but the Multilevel algorithm outperforms all the others.

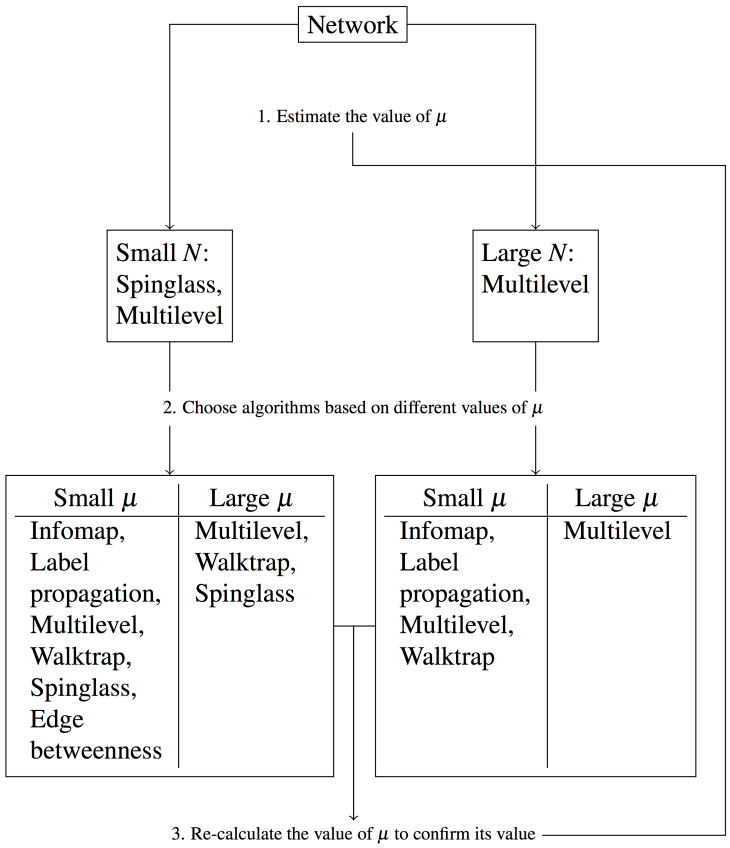
Besides, we illustrate the suggestion for the adaptive use of the methods for community detection process in a simplified flow diagram (see Figure 17). With any given network, one should first employ either Spinglass algorithm or Multilevel algorithm in order to obtain an estimate of the value of the mixing parameter μ. Notice that the former one can only be used for small networks ($N \lessapprox 1000$) due to the prohibitive computing time for larger network sizes. Second, one can choose a suitable method according to the values of N and μ to conduct the community detection such that both the accuracy and the computing time are acceptable. Third,

Figure 16: Recommendation for the choice of adaptable community detection algorithms.



The x-axis is the mixing parameter μ and the y-axis is the number of nodes N. The y-axis is on a *log* scale for better visualisation. The coordinates of certain important points are: $A(0.48, 1000)$, $B(0.6, 1000)$, $C(0.48, 6192)$, $D(0.36, 31948)$, and $E(0.42, 31948)$. In different regions we would like to recommend different algorithms, which are represented by different abbreviations: *IM* is the Infomap algorithm, *LP* is the Label propagation algorithm, *ML* is the Multilevel algorithm, *WT* is the Walktrap algorithm, *SG* is the Spinglass algorithm, and *EB* represents the Edge betweenness algorithm.

Figure 17: Suggestion for the community detection process.



Small networks are those with number of nodes less than 1000, and small μ corresponds to $\mu \lessapprox 0.5$. To be noticed that in the case that $N \geq 1000$ and $\mu \lessapprox 0.5$, Infomap and Multilevel algorithms are no longer suitable choices if $N \geq 6000$.

as we have already shown, in certain situations, there might exist large standard deviations of NMI, i.e., the community detection algorithms are not stable and therefore not reliable. Thus, the value of $\bar{\mu}$ must be recalculated to get an idea of the repeatability of the results and confirm its validity. In some situations, one might need to repeat the detection processes several times or switch to another algorithm to ensure the validity of the community detection results.

Our suggestions have to be applied in conjunction with the concomitant research questions. As a pure application of the recommendations could bias the results. Once a researcher has decided to use a specific community detection algorithm, it is of crucial importance for her to keep in mind the limitations and the expected validity of the output of the community detec-

tion algorithm chosen. It is noteworthy that metadata would be helpful for evaluating network community detection methods and can be used to improve the analysis and understanding of network structure [19, 27]. In real-world networks where metadata is available, researchers should also take into account the research question, the properties of the network, the interpretation and meaning of the communities while choosing the community detection algorithms. Different research questions together with the metadata might lead to different definitions of community, and further change the ground truth of the network.

Compared to previous works on benchmarking community detection algorithms, our study has many obvious advantages: First, we have considered networks which contain a wide spectrum of number of nodes and mixing parameters. Second, the algorithms we have tested are integrated in a cross-platform package which has been widely used in academic research in network science and related fields. Third, we have used the LFR benchmark graphs which have shown more realistic properties than the earlier computer-generated networks such as the GN benchmark.

There are also some limitations in our work: Although the LFR benchmark has generalised the previous GN benchmark by introducing power-law distributions of degree and community size, more realistic properties are still needed. We have mainly focused on testing the effects of the mixing parameter and the number of nodes. Other properties, such as the average degree, the degree distribution exponent, and the community distribution exponent may also play a role in the comparison of algorithms.

In the end, we stress that detecting the community structure of networks is an important issue in network science. For "igraph" package users, we have provided a guideline on choosing the suitable community detection methods. However, based on our results, existing community detection algorithms still need to be improved to better uncover the ground truth of networks.

4.5 Acknowledgements

The authors acknowledge financial support from the URPP Social Networks at University of Zürich. The authors are thankful to the S3IT (Service and Support for Science IT) of the Univer-

sity of Zurich, for providing the support and the computational resources that have contributed to the research results reported in this study, as well as Santo Fortunato for useful comments.

References

[1] Newman, M. E. The structure and function of complex networks. *SIAM Review* **45**, 167–256 (2003).

[2] Boccaletti, S., Latora, V., Moreno, Y., Chavez, M. & Hwang, D.-U. Complex networks: Structure and dynamics. *Physics Reports* **424**, 175–308 (2006).

[3] Girvan, M. & Newman, M. E. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences* **99**, 7821–7826 (2002).

[4] Fortunato, S. Community detection in graphs. *Physics Reports* **486**, 75–174 (2010).

[5] Lancichinetti, A., Fortunato, S. & Radicchi, F. Benchmark graphs for testing community detection algorithms. *Physical Review E* **78**, 046110 (2008).

[6] Lancichinetti, A., Kivelä, M., Saramäki, J. & Fortunato, S. Characterizing the community structure of complex networks. *PloS ONE* **5**, e11976 (2010).

[7] Condon, A. & Karp, R. M. Algorithms for graph partitioning on the planted partition model. *Random Structures and Algorithms* **18**, 116–140 (2001).

[8] Danon, L., Díaz-Guilera, A., Duch, J. & Arenas, A. Comparing community structure identification. *Journal of Statistical Mechanics: Theory and Experiment* **2005**, P09008 (2005).

[9] Barabási, A.-L. & Albert, R. Emergence of scaling in random networks. *Science* **286**, 509–512 (1999).

[10] Palla, G., Derényi, I., Farkas, I. & Vicsek, T. Uncovering the overlapping community structure of complex networks in nature and society. *Nature* **435**, 814–818 (2005).

[11] Guimerà, R., Danon, L., Díaz-Guilera, A., Giralt, F. & Arenas, A. Self-similar community structure in a network of human interactions. *Physical Review E* **68**, 065103 (2003).

[12] Clauset, A., Newman, M. E. & Moore, C. Finding community structure in very large networks. *Physical Review E* **70**, 066111 (2004).

[13] Lancichinetti, A. & Fortunato, S. Benchmarks for testing community detection algorithms on directed and weighted graphs with overlapping communities. *Physical Review E* **80**, 016118 (2009).

[14] Orman, G. K. & Labatut, V. A comparison of community detection algorithms on artificial networks. In *Discovery Science*, 242–256 (Springer, 2009).

[15] Lancichinetti, A. & Fortunato, S. Community detection algorithms: a comparative analysis. *Physical Review E* **80**, 056117 (2009).

[16] Radicchi, F., Castellano, C., Cecconi, F., Loreto, V. & Parisi, D. Defining and identifying communities in networks. *Proceedings of the National Academy of Sciences* **101**, 2658–2663 (2004).

[17] Peel, L. Estimating network parameters for selecting community detection algorithms. In *13th Conference on Information Fusion*, 1–8 (IEEE, 2010).

[18] Hric, D., Darst, R. K. & Fortunato, S. Community detection in networks: Structural communities versus ground truth. *Physical Review E* **90**, 062805 (2014).

[19] Yang, J. & Leskovec, J. Defining and evaluating network communities based on ground-truth. *Knowledge and Information Systems* **42**, 181–213 (2015).

[20] Csardi, G. & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, Complex Systems* 1695. URL `http://igraph.org`.

[21] Vinh, N. X., Epps, J. & Bailey, J. Information theoretic measures for clusterings comparison: Variants, properties, normalization and correction for chance. *The Journal of Machine Learning Research* **11**, 2837–2854 (2010).

[22] Romano, S., Bailey, J., Nguyen, V. & Verspoor, K. Standardized mutual information for clustering comparisons: one step further in adjustment for chance. In *Proceedings of the 31st International Conference on Machine Learning (ICML-14)*, 1143–1151 (2014).

[23] Zhang, P. Evaluating accuracy of community detection using the relative normalized mutual information. *Journal of Statistical Mechanics: Theory and Experiment* **2015**, P11006 (2015).

[24] Papadopoulos, S., Kompatsiaris, Y., Vakali, A. & Spyridonos, P. Community detection in social media. *Data Mining and Knowledge Discovery* **24**, 515–554 (2012).

[25] Blondel, V. D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment* **2008**, P10008 (2008).

[26] Fortunato, S. & Barthélemy, M. Resolution limit in community detection. *Proceedings of the National Academy of Sciences* **104**, 36–41 (2007).

[27] Newman, M. & Clauset, A. Structure and inference in annotated networks. *arXiv preprint arXiv:1507.04001* (2015).

[28] Zachary, W. W. An information flow model for conflict and fission in small groups. *Journal of Anthropological Research* 452–473 (1977).

[29] Danon, L., Díaz-Guilera, A. & Arenas, A. The effect of size heterogeneity on community identification in complex networks. *Journal of Statistical Mechanics: Theory and Experiment* **2006**, P11010 (2006).

[30] Molloy, M. & Reed, B. A critical point for random graphs with a given degree sequence. *Random Structures and Algorithms* **6**, 161–180 (1995).

[31] Bagrow, J. P. Evaluating local community methods in networks. *Journal of Statistical Mechanics: Theory and Experiment* **2008**, P05001 (2008).

[32] Poncela, J., Gómez-Gardeñes, J., Floría, L. M., Sánchez, A. & Moreno, Y. Complex cooperative networks from evolutionary preferential attachment. *PLoS ONE* **3**, e2449 (2008).

[33] Orman, G. K. & Labatut, V. The effect of network realism on community detection algorithms. In *International Conference on Advances in Social Networks Analysis and Mining*, 301–305 (IEEE, 2010).

[34] Freeman, L. C. Centrality in social networks conceptual clarification. *Social Networks* **1**, 215–239 (1979).

[35] Rosvall, M. & Bergstrom, C. T. An information-theoretic framework for resolving community structure in complex networks. *Proceedings of the National Academy of Sciences* **104**, 7327–7331 (2007).

[36] Rosvall, M., Axelsson, D. & Bergstrom, C. T. The map equation. *The European Physical Journal Special Topics* **178**, 13–23 (2010).

[37] Mukherjee, A., Choudhury, M., Peruani, F., Ganguly, N. & Mitra, B. *Dynamics On and Of Complex Networks, Volume 2: Applications to Time-Varying Dynamical Systems* (Springer Science & Business Media, 2013).

[38] Raghavan, U. N., Albert, R. & Kumara, S. Near linear time algorithm to detect community structures in large-scale networks. *Physical Review E* **76**, 036106 (2007).

[39] Newman, M. E. Finding community structure in networks using the eigenvectors of matrices. *Physical Review E* **74**, 036104 (2006).

[40] Xie, J. & Szymanski, B. K. Community detection using a neighborhood strength driven label propagation algorithm. In *Network Science Workshop*, 188–195 (IEEE, 2011).

[41] Reichardt, J. & Bornholdt, S. Statistical mechanics of community detection. *Physical Review E* **74**, 016110 (2006).

[42] Wu, F.-Y. The Potts model. *Reviews of Modern Physics* **54**, 235 (1982).

[43] Kirkpatrick, S. Optimization by simulated annealing: Quantitative studies. *Journal of Statistical Physics* **34**, 975–986 (1984).

[44] Traag, V. & Bruggeman, J. Community detection in networks with positive and negative links. *Physical Review E* **80**, 036115 (2009).

[45] Dahlin, J. & Svenson, P. Ensemble approaches for improving community detection methods. *arXiv:1309.0242 [physics.soc-ph]* (2013).

[46] Pons, P. & Latapy, M. Computing communities in large networks using random walks. In *Computer and Information Sciences-ISCIS 2005*, 284–293 (Springer, 2005).

5 Summary and Outlook

This dissertation has looked into the topic of unethical customer behavior.

Without loss of generality, we discuss the causes of unethical behavior in Chapter 3.3 from four different perspectives, i.e. the standard economic perspective, the psychological perspective, the behavior economic perspective, and the neuroscience perspective. We further discuss the causes of ordinary unethical behavior, which is the unethical behavior committed by ordinary people intentionally or unintentionally, in Chapter 3.4.

We probe the boundaries of the widely-held view about unethical customer behavior by postulating that not all such behaviors are harmful, and further develop a theoretical framework of retailer response to unethical customer behavior in Chapter 2.2. By applying the vector autoregressive models on a longitudinal data from a Swiss online retailer in Chapter 2.3 $\sim$ 2.5, we empirically prove that over longer time periods, the consequences of unethical customer behavior could be positive to both other customers and the retailers.

Moreover, we briefly review existing statistical techniques in fraud detection in Chapter 3.5 and highlight the application of network science in fraud detection in Chapter 3.6. Eight state-of-the-art community detection algorithms, which are the core of "community-based" anomaly detection techniques, have been examined in Chapter 4. We have compared their performances in terms of accuracy and computing time on the Lancichinetti-Fortunato-Radicchi benchmark graphs in Chapter 4.3.

The implications of our studies on the causes and consequences of unethical customer behavior are provided in Chapter 2.7 and 3.7, and the implications of our study on the comparison of community detection algorithms can be found in Chapter 4.4.

Research to come can be in different areas. For instance, how should retailers design/change their user policies to lead to the best possible outcomes for both customers and retailers? Is there a way to integrate the mechanisms of ordinary unethical behavior into the "new fraud triangle" framework? Can we improve existing community detection algorithms or design new methods to better uncover the structure of networks and further improve the performance of

"community-based" anomaly detection techniques?

Nevertheless, we have to be aware of the ethical use of our finding in marketing as well as in sociological perspectives. It's indeed a challenge for companies and individuals to deal with unethical customer behavior, especially in the cases that the consequences of the unethical behavior are not negative.

A Supplementary Information of Chapter 2

A.1 Identifying fraudulent users on the online shopping site

The cookie is typically a tagged string of text that contains data about the user's visit to the website. It flows between a user's computer or mobile device and a web server. If cookie caching has been enabled on the client browser in the user's computer/mobile device, the client browser will store the cookie in the hard drive. For our data source, disabling cookies may lead to the situation that the online platform is not fully functional. The information generated by the cookie about the use of the site includes: browser type and version, operating system, referrer URL (the previously visited page), host name of the accessing computer (IP Address), and time of server request. When registering for a user account on the site, the site requires provision of personal data for successfully delivery of purchased products. This information includes the customer's first name, the last name, the address, the date of birth, and the email address. Together with the data of the transactions with other accounts, the pseudonymous user profiles are created. Cookies in web browsers are also widely used to manage login. A server time stamp may be inserted into cookies indicating when they were created. This information from the cookie can then be compared with other selected information about the user, device and/or account to detect suspicious activities.

The retailer's fraudulent detection procedure contains multiple layers as described in the paper: the account check, the device check, the customer activity check, and most importantly, the manual check [2]. Every account is checked based on their account information, followed by the device information, and after that the historical behavior. The manual check, i.e., a phone call by the retailer's customer service representative, is the last layer in the fraudulent account detection process.

A.2 VAR Model Specification

This appendix provides further information on the (1) lag-order selection, and (2) stability test.

Lag-order selection

Many selection-order statistics are available to select the lag-order of VAR model. In this study we compute four information criteria (FPE, AIC, HQIC, SBIC). The maximum lag order has been set to 30. Two different versions of the information criteria have been applied [5]. Table 22 shows part of the statistics. In the end we set the lag order equals to 2 based on the selection of AIC and FPE.

Table 22: Selection-order statistics. FPE, AIC, HQIC and SBIC are included.

Selection-order criteria (lutstats)				
Sample: 31 - 506, but with a gap. Number of obs: 428				
Lag	**FPE**	**AIC**	**HQIC**	**SBIC**
0	1.00E+46	105.48	105.48	105.48
1	1.30E+45	103.387	103.627*	103.994*
2	1.20E+45*	103.324*	103.804	104.538
3	1.40E+45	103.437	104.156	105.258
Selection-order criteria				
Sample: 31 - 506, but with a gap. Number of obs: 428				
Lag	**FPE**	**AIC**	**HQIC**	**SBIC**
0	1.00E+46	128.669	129.058	129.655
1	1.30E+45	126.576	127.205*	128.169*
2	1.20E+45*	126.513*	127.382	128.714
3	1.40E+45	126.626	127.735	129.433

Two different methods are applied, Lütkepohl's version of information criteria, and the normal version. Endogenous variables: n_PT_t, n_ST_t, n_L_t, n_T_t, f_PT_t, f_ST_t, f_L_t, and f_T_t. Exogenous variables: $adspending_t$, $season1_t$ to $season3_t$, $dofWk1_t$ to $dofWk6_t$, $\#days_t$, and $log(\#days)_t$.

Stability of VAR model

To ensure that the VAR model is stable, we check the eigenvalue stability condition after estimating the parameters of VAR. If any of them exceeds 1 in modulus, some variables in the model may require differencing. Figure 18 presents the results. As all the eigenvalues lie inside the unit circle, VAR satisfies stability condition. The estimates of VAR model are stable and no differencing is necessary.

Figure 18: Stability of VAR model. Eigenvalue stability condition of the estimates of the VAR model has been checked.

Eigenvalue stability condition

Eigenvalue	Modulus
.8581831	.858183
.6967751 + .062955*i*	.699613
.6967751 - .062955*i*	.699613
.6025199	.60252
.5261194 + .2112218*i*	.566936
.5261194 - .2112218*i*	.566936
.4973094	.497309
-.3267023 + .05143208*i*	.330726
-.3267023 - .05143208*i*	.330726
-.2838228 + .1203406*i*	.308281
-.2838228 - .1203406*i*	.308281
-.1595545 + .2368961*i*	.285618
-.1595545 - .2368961*i*	.285618
.03354141 + .1267543*i*	.131117
.03354141 - .1267543*i*	.131117
.02840038	.0284

A.3 Comparison of alternative VAR models

We compare the in-sample and out-of-sample fit of the VAR model with two alternative models in terms of root-mean-square error (RMSE) and median absolute error (MAE). We reserved the last 123 daily observations for a holdout test and re-estimated all models on the first 365 daily observations. Taken the normal accounts' revenue as an example: First, an autoregressive model (AR) captures the immediate effects of the other variables on normal accounts' revenue and the dynamic effect of past normal accounts' revenue on current normal accounts' revenue. Second, an autoregressive-distributed lag model (ARDL) will add dynamic effect of each other variable, without accounting for endogeneity and indirect effects among these variables.

AR model

The AR specification in Equation 5 relates normal accounts' login activity to all the other variables related to fraudulent and normal accounts and controls for advertisement spending, time

trend, seasonality, and lags of the dependent variable [1]. This equation includes the same variables as the VAR model (see Figure 3, Formula of the VAR model, for details), J is the number of lags of the dependent variable (i.e. the term "autoregressive").

ARDL model

In the AR model, we only capture the immediate effects of the other variables on normal accounts' revenue. To include dynamic effect specific to the other variables, we add lags of those variables and make the following ARDL model [4]. This model includes the same variables as the VAR model, J is the number of lags of the dependent variable. L, M, N, P, Q, R and S are the number of lags for the predictor variables number of logins of fraudulent accounts, number of proposed transactions of fraudulent accounts, number of successful transactions of fraudulent accounts, amount of revenue of fraudulent accounts, number of transactions proposed by normal accounts, number of successful transactions made by normal accounts, and number of logins of normal accounts, respectively (See Equation 6 for details).

The comparison results in Table 23 have shown that: 1) In the out-of-sample comparison, the VAR model produces much lower values for RMSE and MAE, indicating superior fit to the data; and 2) in the in-sample comparison, AR model is slightly better than the other two models in this case. It is mainly due to the fact that only the AR model has captured the immediate effects of the other variables on normal accounts' revenue while the ARDL and VAR models are focusing on the dynamic effects.

$$\begin{aligned} n\_T_t = {} & \beta_1 f\_L_t + \beta_2 f\_PT_t + \beta_3 f\_ST_t + \beta_4 f\_T_t + \beta_5 n\_PT_t + \beta_6 n\_ST_t + \beta_7 n\_L_t \\ & + \beta_8 adspending_t + \beta_9 season1_t + \beta_{10} season2_t + \beta_{11} season3_t \\ & + \beta_{12} dofWK1_t + \beta_{13} dofWK2_t + \beta_{14} dofWK3_t + \beta_{15} dofWK4_t \\ & + \beta_{16} dofWK5_t + \beta_{17} dofWK6_t + \beta_{18} time\ trends + intercept \\ & + \sum_{j=1}^{J} \gamma_j n\_T_{t-j} + \varepsilon_t \end{aligned} \tag{5}$$

$$
\begin{aligned}
n\_T_t = & \sum_{l=1}^{L} \beta_{1,l} f\_L_{t-l} + \sum_{m=1}^{M} \beta_{2,m} f\_PT_{t-m} + \sum_{n=1}^{N} \beta_{3,n} f\_ST_{t-n} + \sum_{p=1}^{P} \beta_{4,p} f\_T_{t-p} \\
& + \sum_{q=1}^{Q} \beta_{5,q} n\_PT_{t-q} + \sum_{r=1}^{R} \beta_{6,r} n\_ST_{t-r} + \sum_{s=1}^{S} \beta_{7,s} n\_L_{t-s} + \beta_8 adspending_t \\
& + \beta_9 season1_t + \beta_{10} season2_t + \beta_{11} season3_t + \beta_{12} dofWK1_t + \beta_{13} dofWK2_t \\
& + \beta_{14} dofWK3_t + \beta_{15} dofWK4_t + \beta_{16} dofWK5_t + \beta_{17} dofWK6_t + \beta_{18} time\ trends \\
& + intercept + \sum_{j=1}^{J} \gamma_j n\_T_{t-j} + \varepsilon_t
\end{aligned}
\tag{6}
$$

Table 23: Comparison of VAR with AR and ARDL in terms of root-mean-square error (RMSE) and median absolute error (MAE).

		In-sample (1-365)		Out-of-sample (366-488)	
Variable	Models	RMSE	MAE	RMSE	MAE
n_T	VAR	77934.34	54210.83	58908.26	42466.5
	AR	71999.5	51087.97	106921.37	92915.62
	ARDL	77534.38	53966.82	73406.29	58260.5

A.4 Robustness check based on different definitions of fraudulent accounts

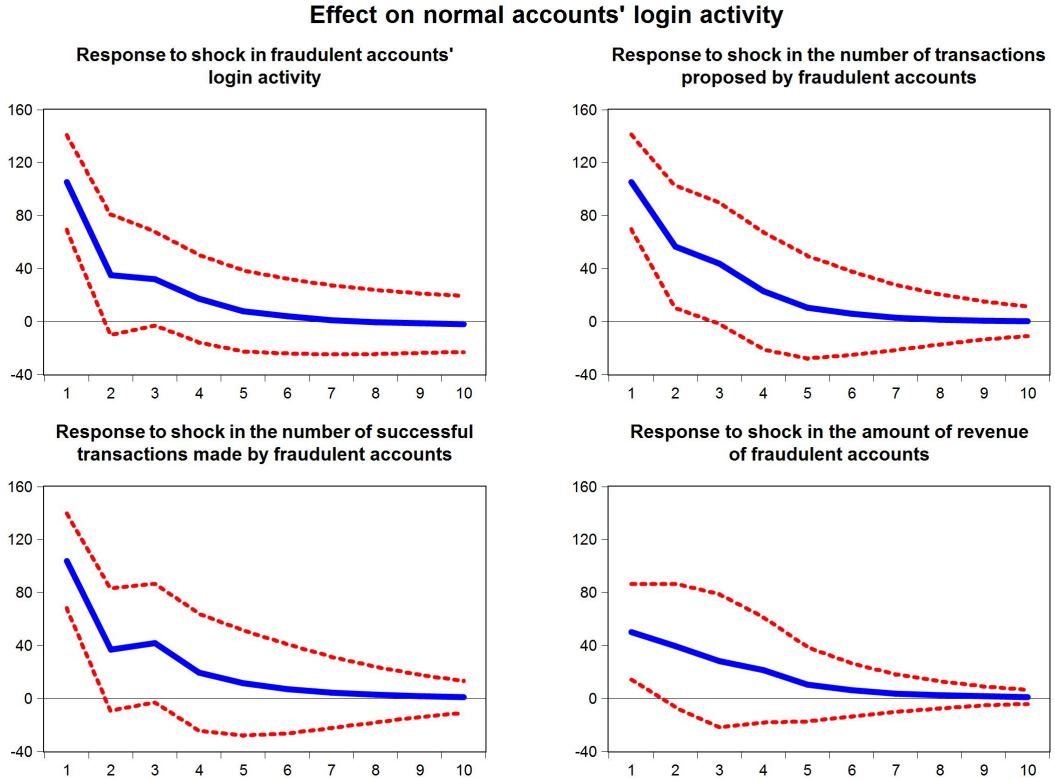
We rerun the VAR model by using different definitions of fraudulent accounts. The aim of the robustness check is to better validate our results.

Fraudulent accounts as those which have received the email and the in-game warning message

In this part, we used the list of fraudulent accounts that received both the email and the in-game warning message. There are 4,345 unique accounts on this list.

In the figure 19, the blue curve shows the impact and the red dotted curves represent the one standard deviation confidence interval for the impact. The x-axis is the number of days. The y-axis is the number of login times. An unexpected positive shock on fraudulent accounts' login times yields an increase of 105 login times on normal accounts' login activity. The corresponding elasticity is 0.0535. This positive effect decreases over time, disappearing in 2 days. An unexpected positive shock on fraudulent accounts' number of proposed transactions yields an increase of 105 login times on normal accounts' login activity. The corresponding elasticity is 0.0415. This positive effect goes to zero after 3 days. An unexpected positive shock on fraudulent accounts' number of successful transactions yields an increase of 104 login times on normal accounts' login activity. The corresponding elasticity is 0.0493. This positive effect becomes insignificant in 2 days. An unexpected positive shock on fraudulent accounts' amount of revenue does not yield any significant response on normal accounts' login activity.

Figure 19: The plots of the generalized impulse response functions for the impact of an unexpected change in fraudulent accounts' variables on normal accounts' login activity.

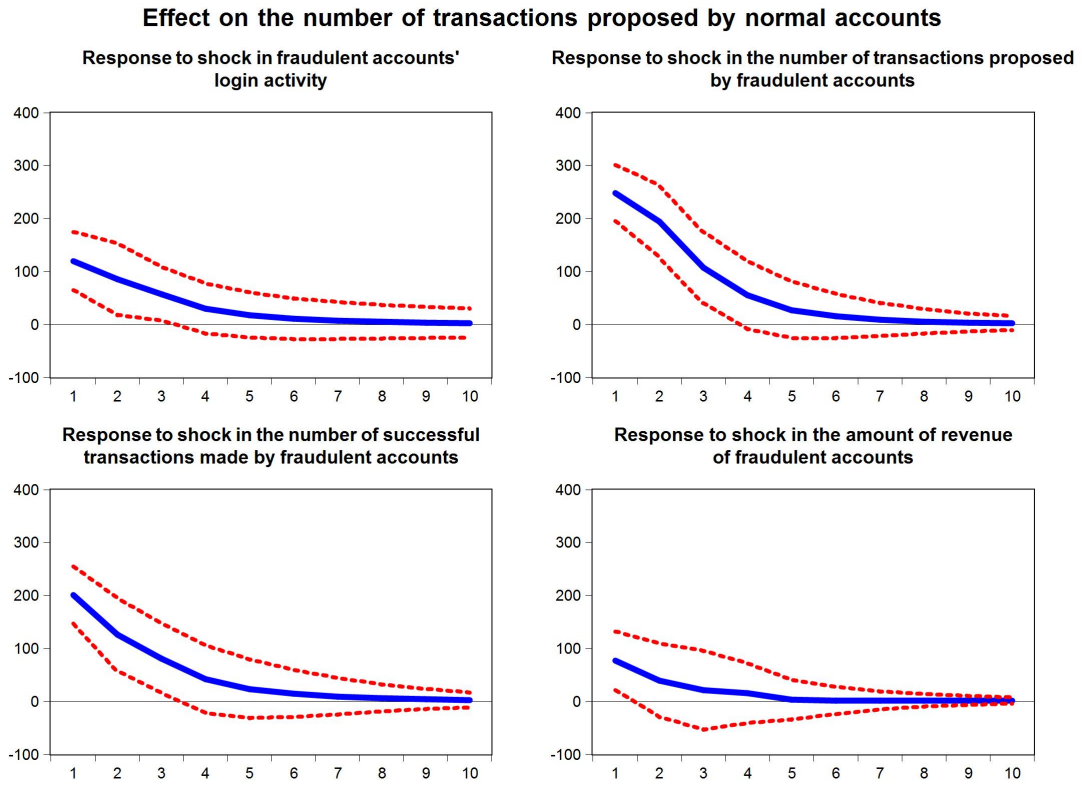


In the figure 20, the blue curve shows the impact and the red dotted curves represent the one standard deviation confidence interval for the impact. The x-axis is the number of days. The y-axis is the number of proposed transaction times. An unexpected positive shock on fraudulent accounts' login times yields an increase of 120 transaction times on normal accounts' number of proposed transactions. The corresponding elasticity is 0.0793. This positive effect decreases over time and disappear after 3 days. An unexpected positive shock on fraudulent accounts' number of proposed transactions yields an increase of 248 units on normal accounts' number of proposed transactions. The corresponding elasticity is 0.1274. This positive effect goes to zero after 4 days. An unexpected positive shock on fraudulent accounts' number of successful transactions leads to 201 more proposed transactions of normal account. The corresponding elasticity is 0.1236. This positive effect lasts for 3 days before it becomes insignificant. An unexpected positive shock on fraudulent accounts' amount of revenue does not yield any significant response on the number of normal accounts' proposed transactions.

In the figure 21, the blue curve shows the impact and the red dotted curves represent the one standard deviation confidence interval for the impact. The x-axis is the number of days. The y-axis is the number of successful transaction times. An unexpected positive shock on fraudulent accounts' login times yields an increase of 35 transaction times on normal accounts' number of successful transactions. The corresponding elasticity is 0.0873. This positive effect decreases over time and disappears after 5 days. An unexpected positive shock on fraudulent accounts' number of proposed transactions yields an increase of 60 units on normal accounts' number of successful transactions. The corresponding elasticity is 0.1162. This positive effect goes to zero after 4 days. An unexpected positive shock on fraudulent accounts' number of successful transactions leads to 61 more successful transactions of the normal account. The corresponding elasticity is 0.1415. This positive effect lasts for 4 days before it becomes insignificant. An unexpected positive shock on fraudulent accounts' amount of revenue does not yield any significant response on the number of normal accounts' successful transactions.

In the figure 22, the blue curve shows the impact and the red dotted curves represent the one standard deviation confidence interval for the impact. The x-axis is the number of days.

Figure 20: The plots of the generalized impulse response functions for the impact of an unexpected change in fraudulent accounts' variables on the number of transactions proposed by normal accounts.

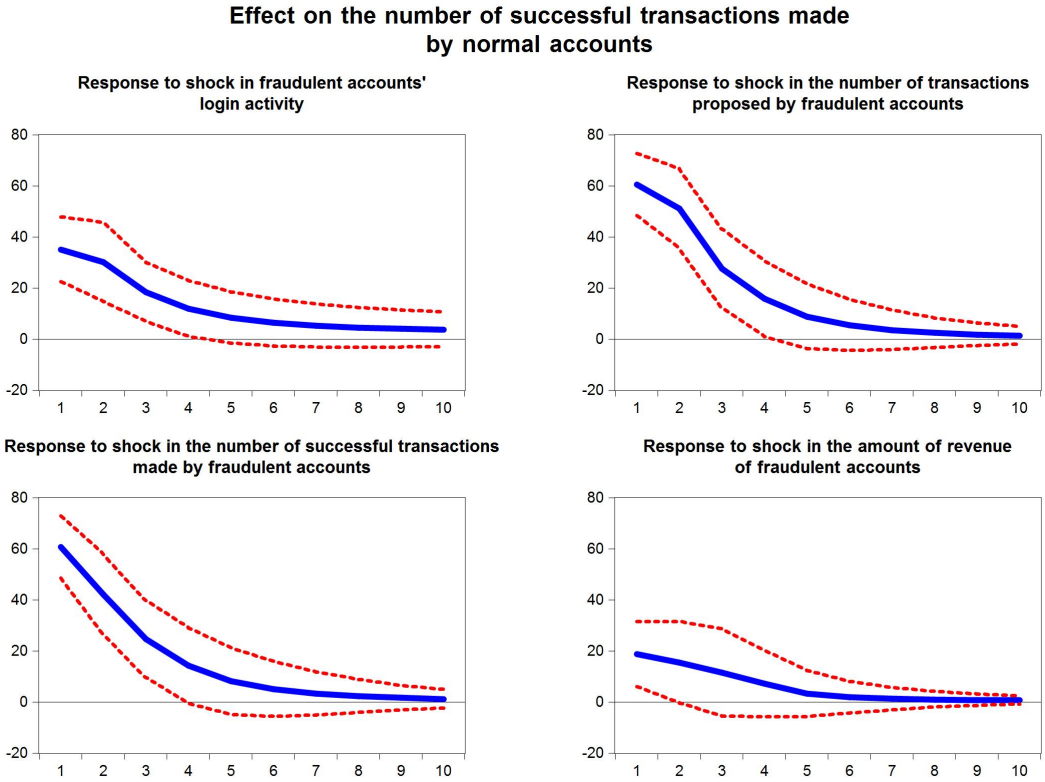


The *y*-axis is the amount of revenue in 0.01 CHF. An unexpected positive shock on fraudulent accounts' login times, number of proposed transactions, and number of successful transactions does not yield any significant response on normal accounts' revenue. An unexpected positive shock on fraudulent accounts' amount of revenue yields an increase of 24,067 units in the amount of revenue of normal accounts. The corresponding elasticity is 0.1475. This positive effect becomes insignificant in about 2 days.

Fraudulent accounts as those which have received the email, the in-game warning message, plus the termination message

In this part, we included only fraudulent accounts that received all three treatments – the email, the in-game warning, and the termination message. There are 3,749 unique accounts on this

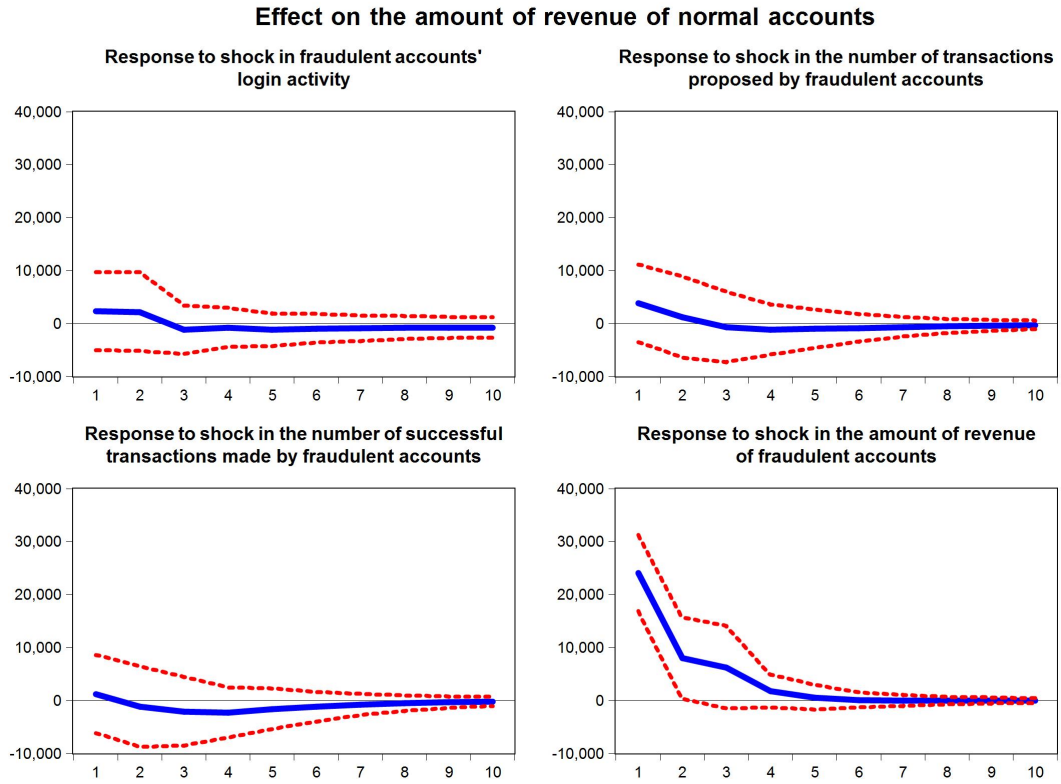
Figure 21: The plots of the generalized impulse response functions for the impact of an unexpected change in fraudulent accounts' variables on the number of successful transaction made by normal accounts.



list.

In the figure 23, the blue curve shows the impact and the red dotted curves represent the one standard deviation confidence interval for the impact. The x-axis is the number of days. The y-axis is the number of login times. An unexpected positive shock on fraudulent accounts' login times yields an increase of 94 login times on normal accounts' login activity. The corresponding elasticity is 0.0404. This positive effect decreases over time and disappears in 2 days. An unexpected positive shock on fraudulent accounts' number of proposed transactions yields an increase of 105 login times on normal accounts' login activity. The corresponding elasticity is 0.036. This positive effect goes to zero after 3 days. An unexpected positive shock on fraudulent accounts' number of successful transactions yields an increase of 103 login times on normal accounts' login activity. The corresponding elasticity is 0.0417. This positive effect

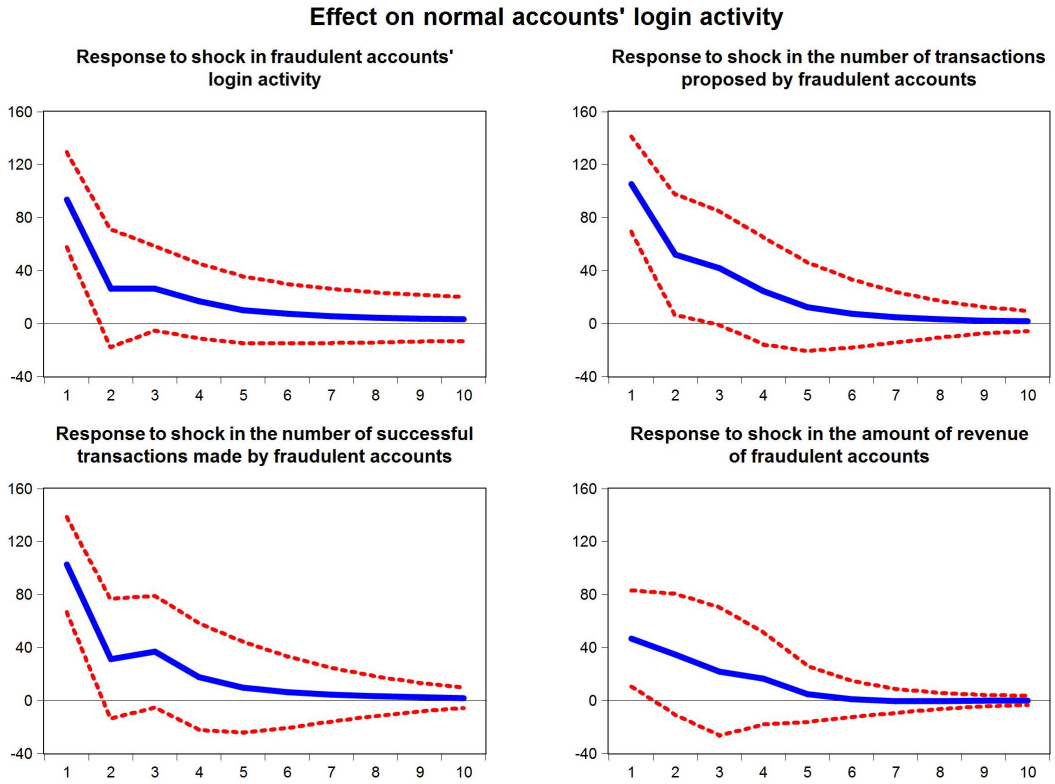
Figure 22: The plots of the generalized impulse response functions for the impact of an unexpected change in fraudulent accounts' variables on the amount of revenue of normal accounts.



becomes insignificant in 2 days. An unexpected positive shock on fraudulent accounts' amount of revenue does not yield any significant response on normal accounts' login activity.

In the figure 24, the blue curve shows the impact and the red dotted curves represent the one standard deviation confidence interval for the impact. The x-axis is the number of days. The y-axis is the number of proposed transaction times. An unexpected positive shock on fraudulent accounts' login times yields an increase of 124 transaction times on normal accounts' number of proposed transactions. The corresponding elasticity is 0.0697. This positive effect decreases over time and disappears after 4 days. An unexpected positive shock on fraudulent accounts' number of proposed transactions yields an increase of 244 units on normal accounts' number of proposed transactions. The corresponding elasticity is 0.1094. This positive effect goes to zero after 4 days. An unexpected positive shock on fraudulent accounts' number of successful

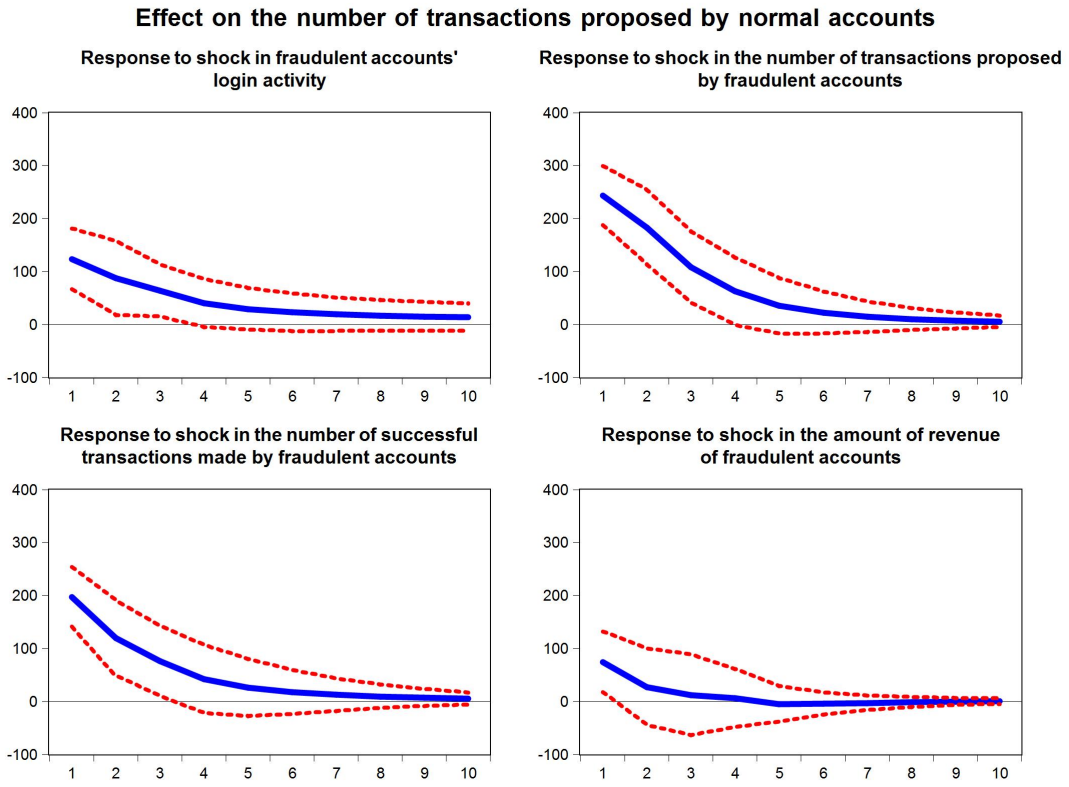
Figure 23: The plots of the generalized impulse response functions for the impact of an unexpected change in fraudulent accounts' variables on normal accounts' login activity.



transactions leads to 197 more proposed transactions of normal account. The corresponding elasticity is 0.1042. This positive effect lasts for 3 days before it becomes insignificant. An unexpected positive shock on fraudulent accounts' amount of revenue does not yield any significant response on the number of normal accounts' proposed transactions.

In the figure 25, the blue curve shows the impact and the red dotted curves represent the one standard deviation confidence interval for the impact. The x-axis is the number of days. The y-axis is the number of successful transaction times. An unexpected positive shock on fraudulent accounts' login times yields an increase of 36 transaction times on normal accounts' number of successful transactions. The corresponding elasticity is 0.0752. This positive effect decreases over time and disappears after 6 days. An unexpected positive shock on fraudulent accounts' number of proposed transactions yields an increase of 61 units on normal accounts' number

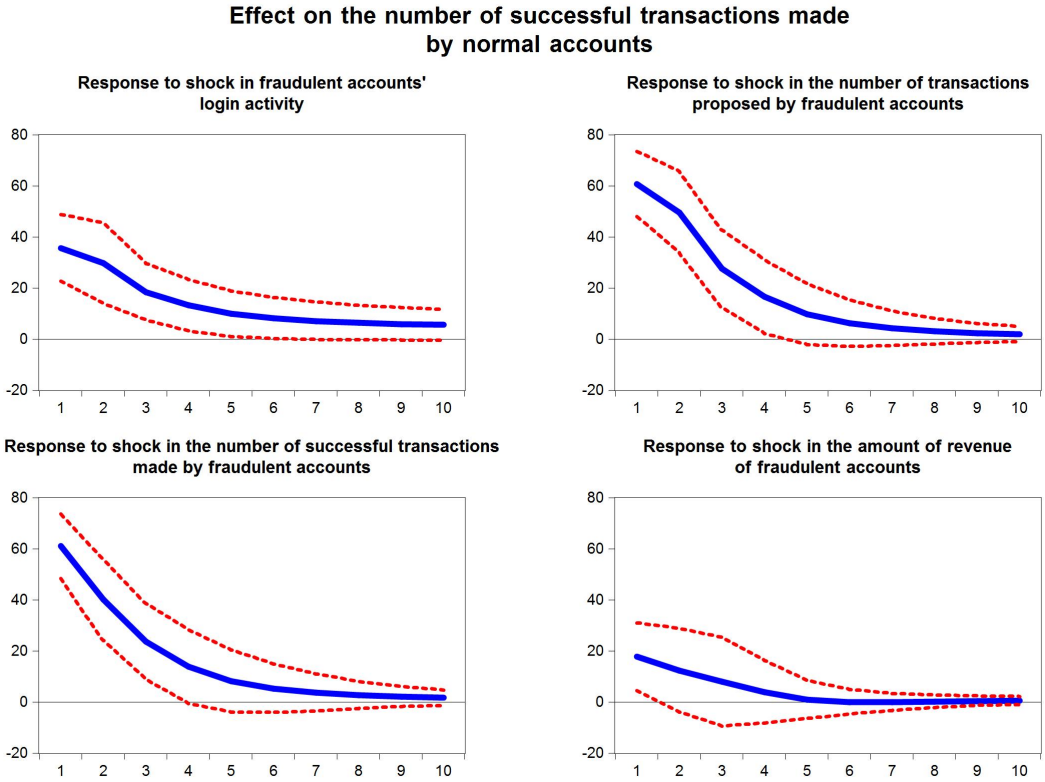
Figure 24: The plots of the generalized impulse response functions for the impact of an unexpected change in fraudulent accounts' variables on the number of transactions proposed by normal accounts.



of successful transactions. The corresponding elasticity is 0.1016. This positive effect goes to zero after 4 days. An unexpected positive shock on fraudulent accounts' number of successful transactions leads to 61 more successful transactions of normal account. The corresponding elasticity is 0.1199. This positive effect lasts for 4 days before it becomes insignificant. An unexpected positive shock on fraudulent accounts' amount of revenue does not yield any significant response on the number of normal accounts' successful transactions.

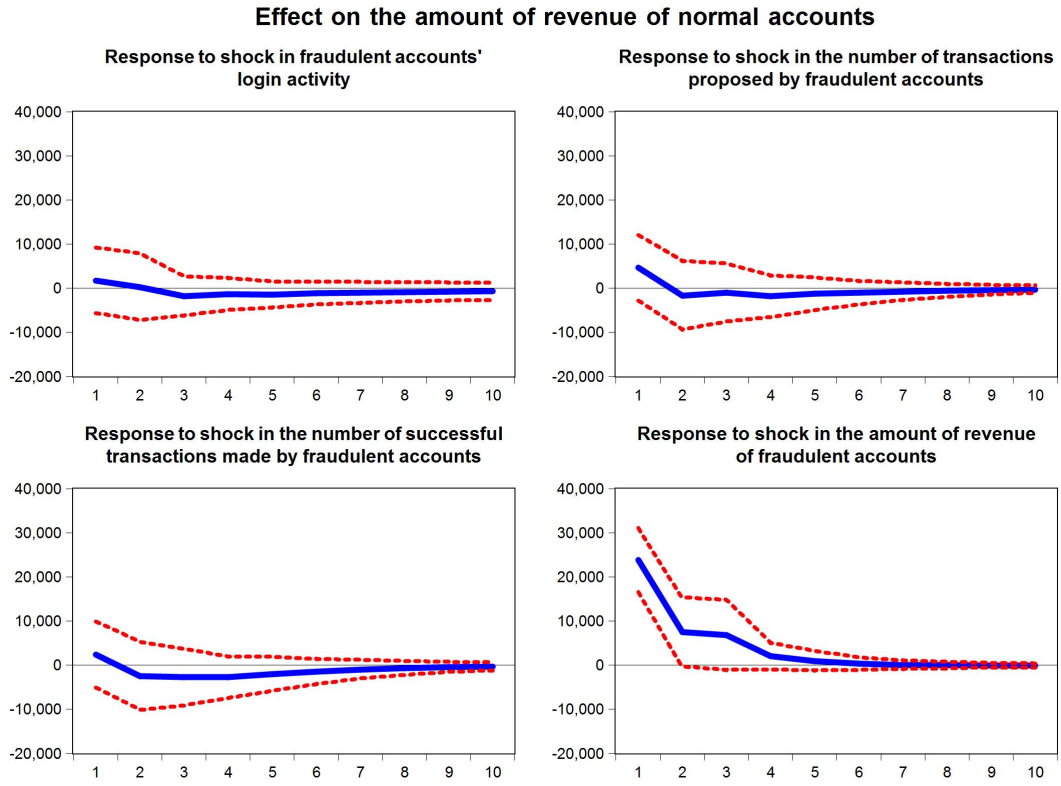
In the figure 26, the blue curve shows the impact and the red dotted curves represent the one standard deviation confidence interval for the impact. The x-axis is the number of days. The y-axis is the amount of revenue in 0.01 CHF. An unexpected positive shock on fraudulent accounts' login times, number of proposed transactions, and number of successful transactions does not yield any significant response on normal accounts' revenue. An unexpected positive

Figure 25: The plots of the generalized impulse response functions for the impact of an unexpected change in fraudulent accounts' variables on the number of successful transaction made by normal accounts.



shock on fraudulent accounts' amount of revenue yields an increase of 23,853 units in the amount of revenue of normal accounts. The corresponding elasticity is 0.1223. This positive effect becomes insignificant in about 2 days.

Figure 26: The plots of the generalized impulse response functions for the impact of an unexpected change in fraudulent accounts' variables on the amount of revenue of normal accounts.



A.5 The Ethics Position Questionnaire (Forsyth 1980)

1. People should make certain that their actions never intentionally harm another even to a small degree.
2. Risks to another should never be tolerated, irrespective of how small the risks might be.
3. The existence of potential harm to others is always wrong, irrespective of the benefits to be gained.
4. One should never psychologically or physically harm another person.
5. One should not perform an action which might in any way threaten the dignity and welfare of another individual.
6. If an action could harm an innocent other, then it should not be done.
7. Deciding whether or not to perform an act by balancing the positive consequences of the act against the negative consequences of the act is immoral.

8. The dignity and welfare of the people should be the most important concern in any society.
9. It is never necessary to sacrifice the welfare of others.
10. Moral behaviors are actions that closely match ideals of the most "perfect" action.
11. There are no ethical principles that are so important that they should be a part of any code of ethics.
12. What is ethical varies from one situation and society to another.
13. Moral standards should be seen as being individualistic; what one person considers to be moral may be judged to be immoral by another person.
14. Different types of morality cannot be compared as to "rightness."
15. Questions of what is ethical for everyone can never be resolved since what is moral or immoral is up to the individual.
16. Moral standards are simply personal rules that indicate how a person should behave, and are not be applied in making judgments of others.
17. Ethical considerations in interpersonal relations are so complex that individuals should be allowed to formulate their own individual codes.
18. Rigidly codifying an ethical position that prevents certain types of actions could stand in the way of better human relations and adjustment.
19. No rule concerning lying can be formulated; whether a lie is permissible or not permissible totally depends upon the situation.
20. Whether a lie is judged to be moral or immoral depends upon the circumstances surrounding the action.

References

[1] Box, G. E. & Jenkins, G. M. Time series models for forecasting and control. *San Francisco* (1970).

[2] Chen, J., Tao, Y., Wang, H. & Chen, T. Big data based fraud risk management at alibaba. *The Journal of Finance and Data Science* **1**, 1–10 (2015).

[3] Forsyth, D. R. A taxonomy of ethical ideologies. *Journal of Personality and Social psychology* **39**, 175 (1980).

[4] Hanssens, D. M., Parsons, L. J. & Schultz, R. L. *Market response models: Econometric and time series analysis*, vol. 12 (Kluwer Academic Publishers, 2001).

[5] Lütkepohl, H. *New introduction to multiple time series analysis* (Berlin: Springer Science & Business Media, 2007).

B Supplementary Information of Chapter 4

Different normalization methods for mutual information

The accuracy of different community detection algorithms can be evaluated by the *normalised mutual information* [1]. As it has been pointed out by Vinh *et al.*, there exist five different normalised versions of the mutual information [2]: I_{joint} ($= \frac{i(\mathscr{P},\bar{\mathscr{P}})}{H(\mathscr{P},\bar{\mathscr{P}})}$), I_{max} ($= \frac{i(\mathscr{P},\bar{\mathscr{P}})}{max\{H(\mathscr{P}),H(\bar{\mathscr{P}})\}}$), I_{sum} ($= \frac{i(\mathscr{P},\bar{\mathscr{P}})}{\frac{1}{2}(H(\mathscr{P})+H(\bar{\mathscr{P}}))}$), I_{sqrt} ($= \frac{i(\mathscr{P},\bar{\mathscr{P}})}{\sqrt{H(\mathscr{P})H(\bar{\mathscr{P}})}}$), and I_{min} ($= \frac{i(\mathscr{P},\bar{\mathscr{P}})}{min\{H(\mathscr{P}),H(\bar{\mathscr{P}})\}}$). Different normalisation methods are sensitive to different partition properties and have different theoretical properties.

In this "Supplementary information", we show the effect of the mixing parameter and network size on all five different NMIs and conclude that the results are similar to each other. In the main text, we report the results of I_{sum} [2], which is consistent with Danon *et al.* [1].

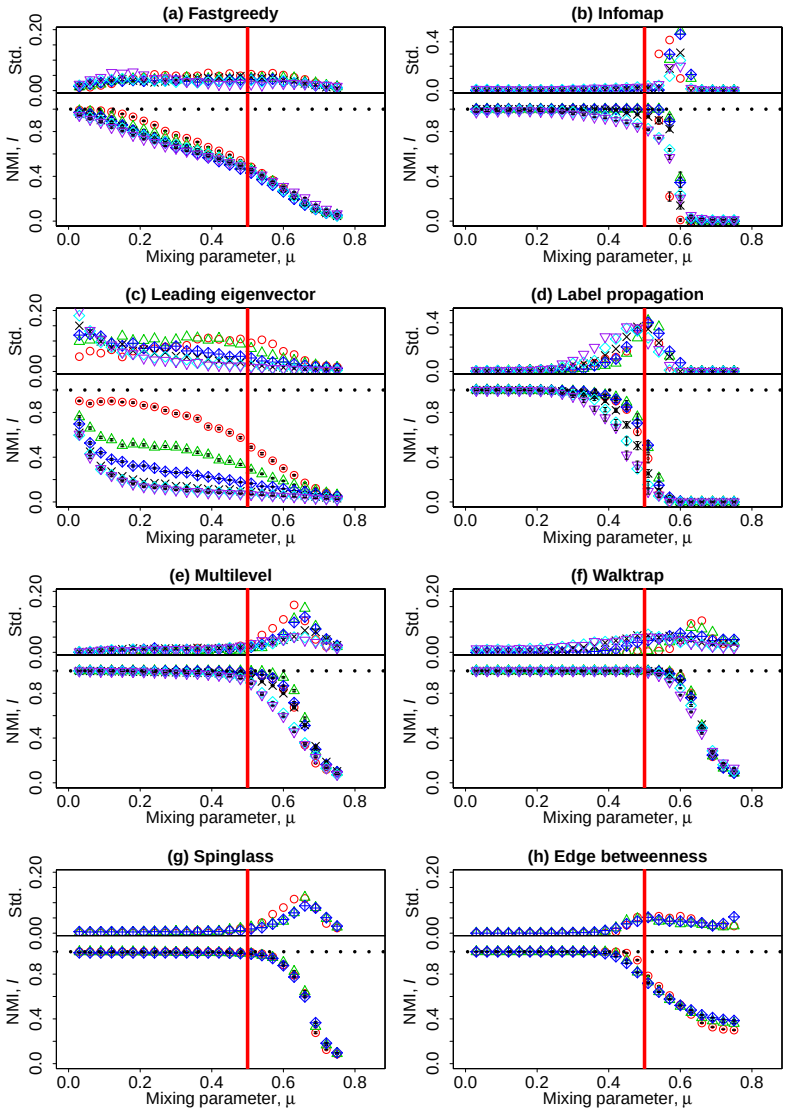
B.1 The role of the network mixing parameter on accuracy

In Figure 27, 28, 29, 30, and 31, we show the effect of the mixing parameter on I_{joint}, I_{max}, I_{sum}, I_{sqrt}, and I_{min}, separately. The detailed explanation of the plot I_{sum} can be found in the main text. Comparing different figures, we conclude that: (1) I_{joint} provides the smallest values and I_{min} provides the largest ones, and (2) all the NMIs display similar patterns.

B.2 The role of network size on accuracy

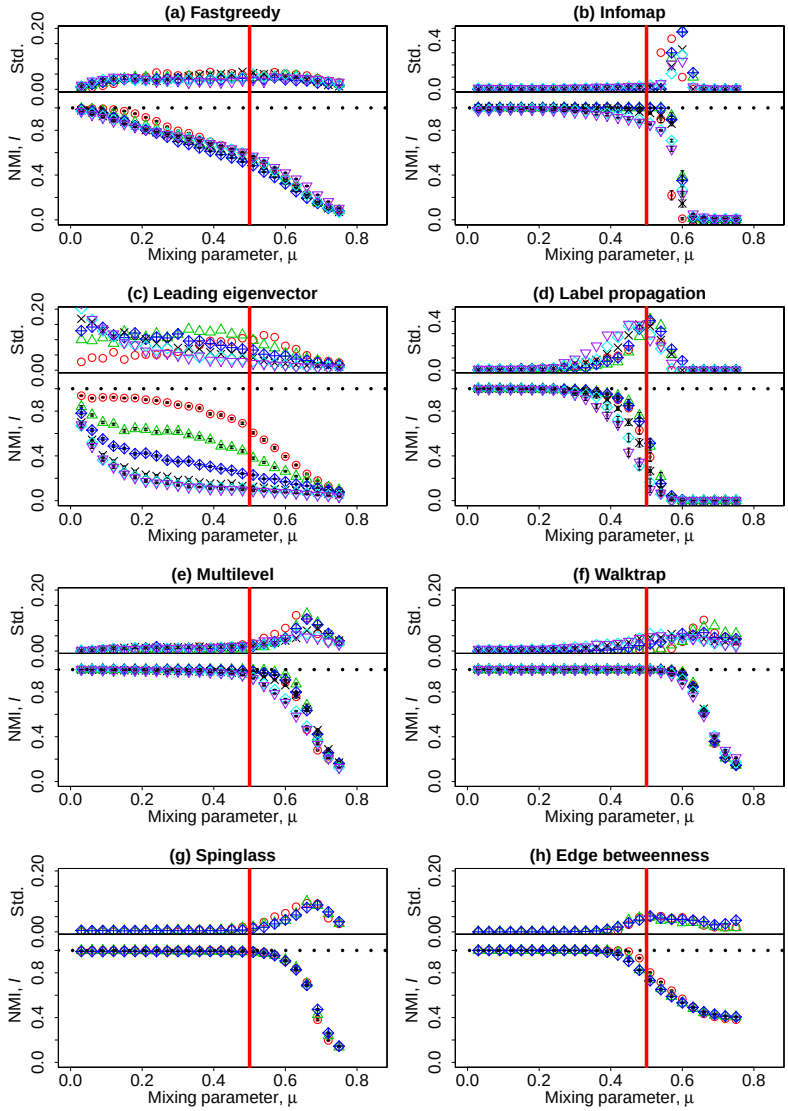
In Figure 32, 33, 34, 35, and 36, we show the effect of the network size on I_{joint}, I_{max}, I_{sum}, I_{sqrt}, and I_{min}, separately. Comparing the different plots we get the same conclusion as before.

Figure 27: (lower row) The mean value of I_{joint} dependent on the mixing parameter μ. (upper row) The standard deviation of I_{joint} dependent on μ.



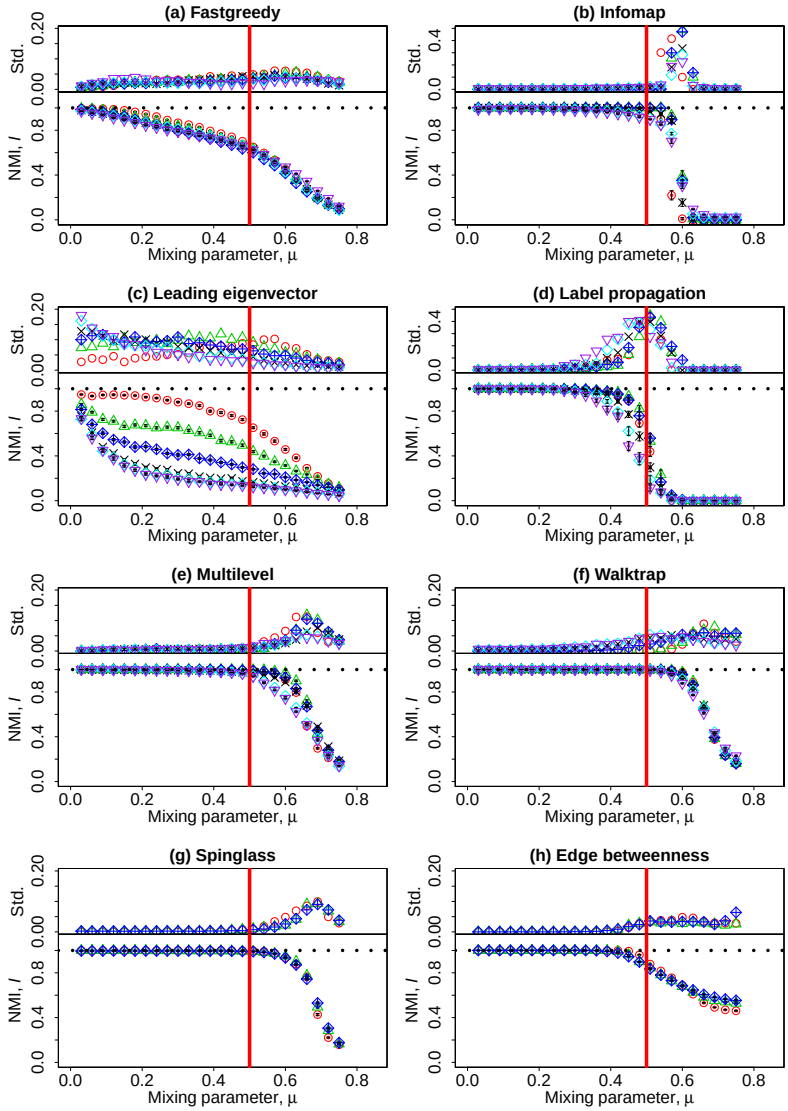
Different colours refer to different number of nodes: red ($N = 233$), green ($N = 482$), blue ($N = 1000$), black ($N = 3583$), cyan ($N = 8916$), and purple ($N = 22186$). Please notice that the vertical axis on the subfigures might have different scale ranges. The vertical red line corresponds to the strong definition of community where $\mu = 0.5$. The horizontal black dotted line corresponds to $I = 1$. The other parameters are described in the main text.

Figure 28: (lower row) The mean value of I_{max} dependent on the mixing parameter μ. (upper row) The standard deviation of I_{max} dependent on μ.



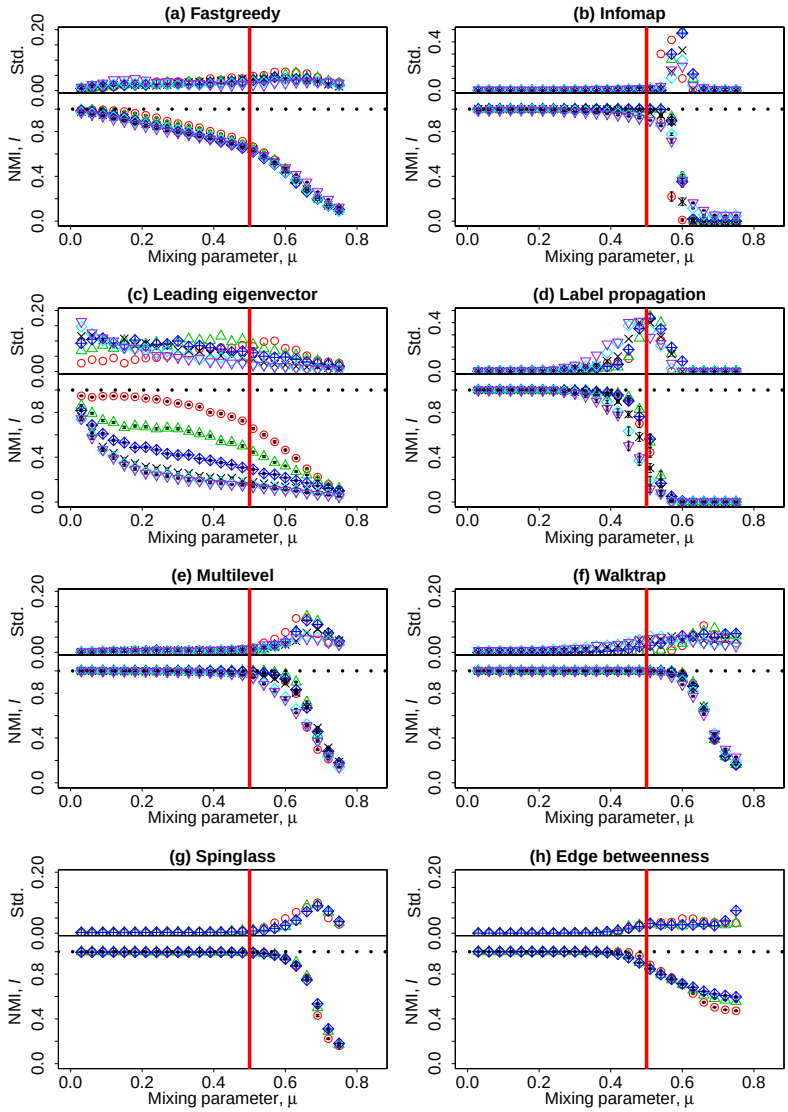
Different colours refer to different number of nodes: red ($N = 233$), green ($N = 482$), blue ($N = 1000$), black ($N = 3583$), cyan ($N = 8916$), and purple ($N = 22186$). Please notice that the vertical axis on the subfigures might have different scale ranges. The vertical red line corresponds to the strong definition of community where $\mu = 0.5$. The horizontal black dotted line corresponds to $I = 1$. The other parameters are described in the main text.

Figure 29: (lower row) The mean value of I_{sum} dependent on the mixing parameter μ. (upper row) The standard deviation of I_{sum} dependent on μ.



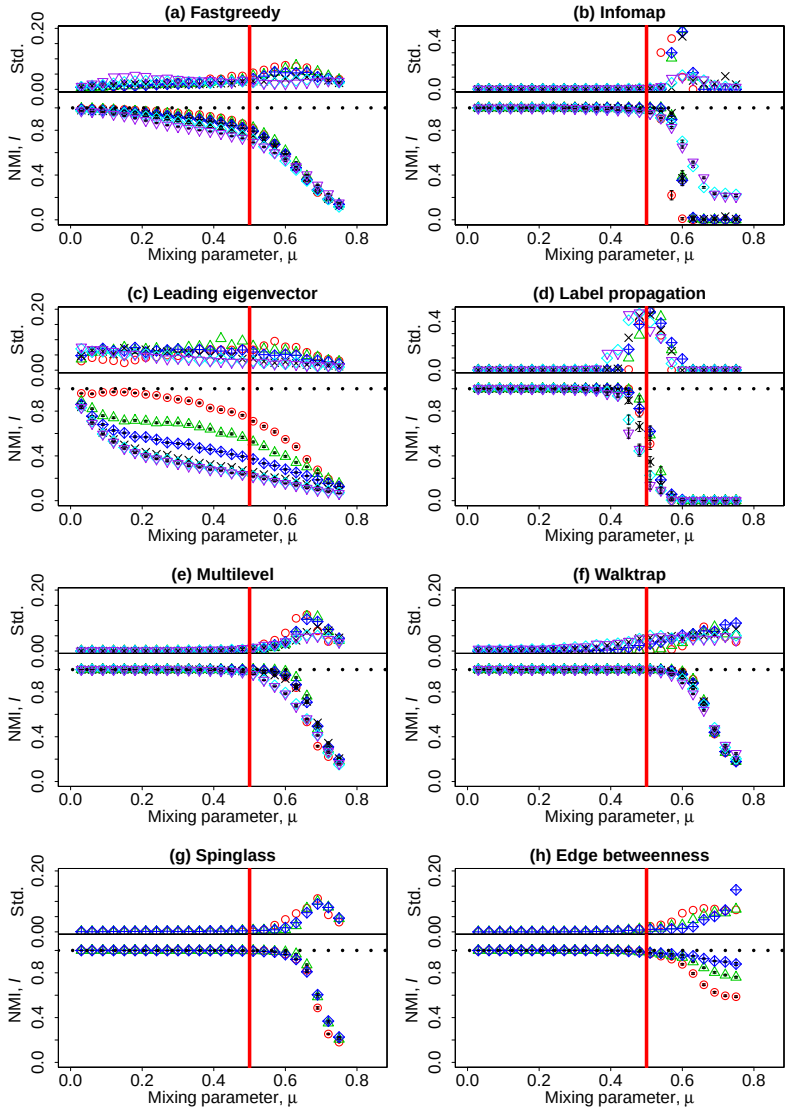
Different colours refer to different number of nodes: red ($N = 233$), green ($N = 482$), blue ($N = 1000$), black ($N = 3583$), cyan ($N = 8916$), and purple ($N = 22186$). Please notice that the vertical axis on the subfigures might have different scale ranges. The vertical red line corresponds to the strong definition of community where $\mu = 0.5$. The horizontal black dotted line corresponds to $I = 1$. The other parameters are described in the main text.

Figure 30: (lower row) The mean value of I_{sqrt} dependent on the mixing parameter μ. (upper row) The standard deviation of I_{sqrt} dependent on μ.



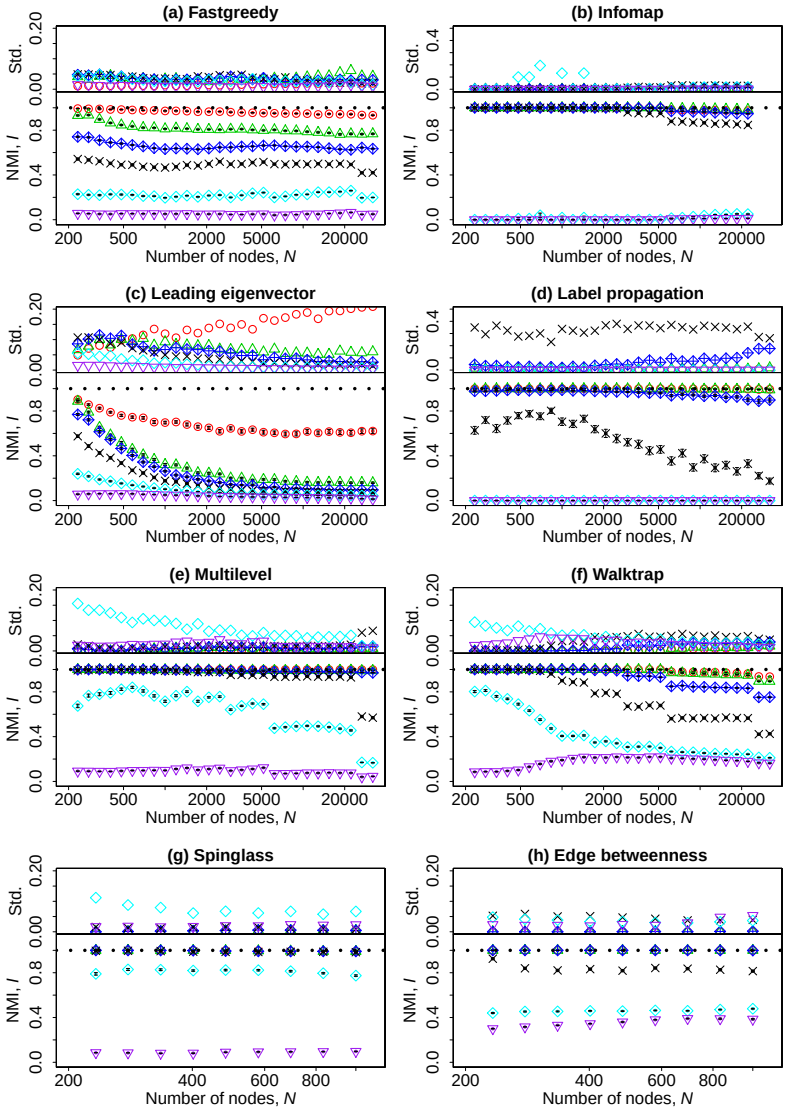
Different colours refer to different number of nodes: red ($N = 233$), green ($N = 482$), blue ($N = 1000$), black ($N = 3583$), cyan ($N = 8916$), and purple ($N = 22186$). Please notice that the vertical axis on the subfigures might have different scale ranges. The vertical red line corresponds to the strong definition of community where $\mu = 0.5$. The horizontal black dotted line corresponds to $I = 1$. The other parameters are described in the main text.

Figure 31: (lower row) The mean value of I_{min} dependent on the mixing parameter μ. (upper row) The standard deviation of I_{min} dependent on μ.



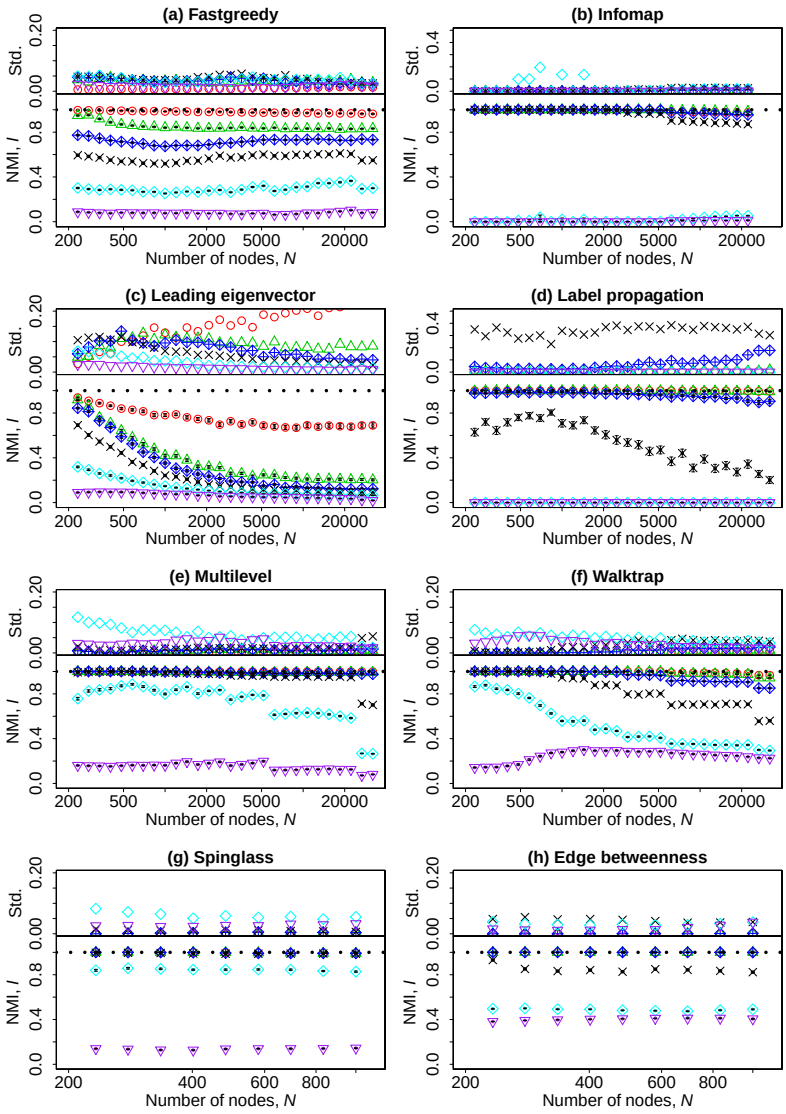
Different colours refer to different number of nodes: red ($N = 233$), green ($N = 482$), blue ($N = 1000$), black ($N = 3583$), cyan ($N = 8916$), and purple ($N = 22186$). Please notice that the vertical axis on the subfigures might have different scale ranges. The vertical red line corresponds to the strong definition of community where $\mu = 0.5$. The horizontal black dotted line corresponds to $I = 1$. The other parameters are described in the main text.

Figure 32: (lower row) The mean value of I_{joint} dependent on the number of nodes N in the benchmark graphs on a *linear-log* scale. (upper row) The standard deviation of I_{joint} dependent on N on a *linear-log* scale.



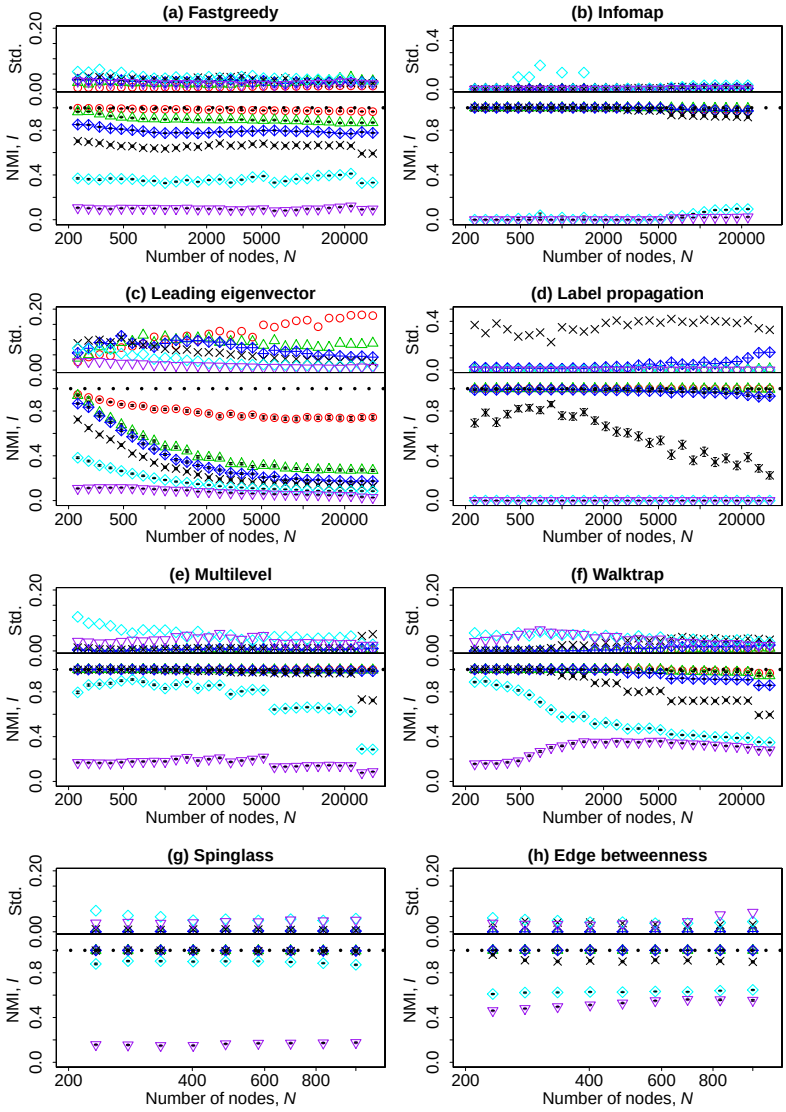
Different colours refer to different values of the mixing parameter: red ($\mu = 0.03$), green ($\mu = 0.18$), blue ($\mu = 0.33$), black ($\mu = 0.48$), cyan ($\mu = 0.63$), and purple ($\mu = 0.75$). Please notice that the vertical axis on the subfigures might have different scale ranges. The horizontal black dotted line corresponds to $I = 1$. Due to the computing speed, Spinglass and Edge betweenness algorithms have been tested only on networks with $N \leq 1000$, and Infomap algorithm has been tested on networks with $N \leq 22186$. The other parameters are described in the main text.

Figure 33: (lower row) The mean value of I_{max} dependent on the number of nodes N in the benchmark graphs on a *linear-log* scale. (upper row) The standard deviation of I_{max} dependent on N on a *linear-log* scale.



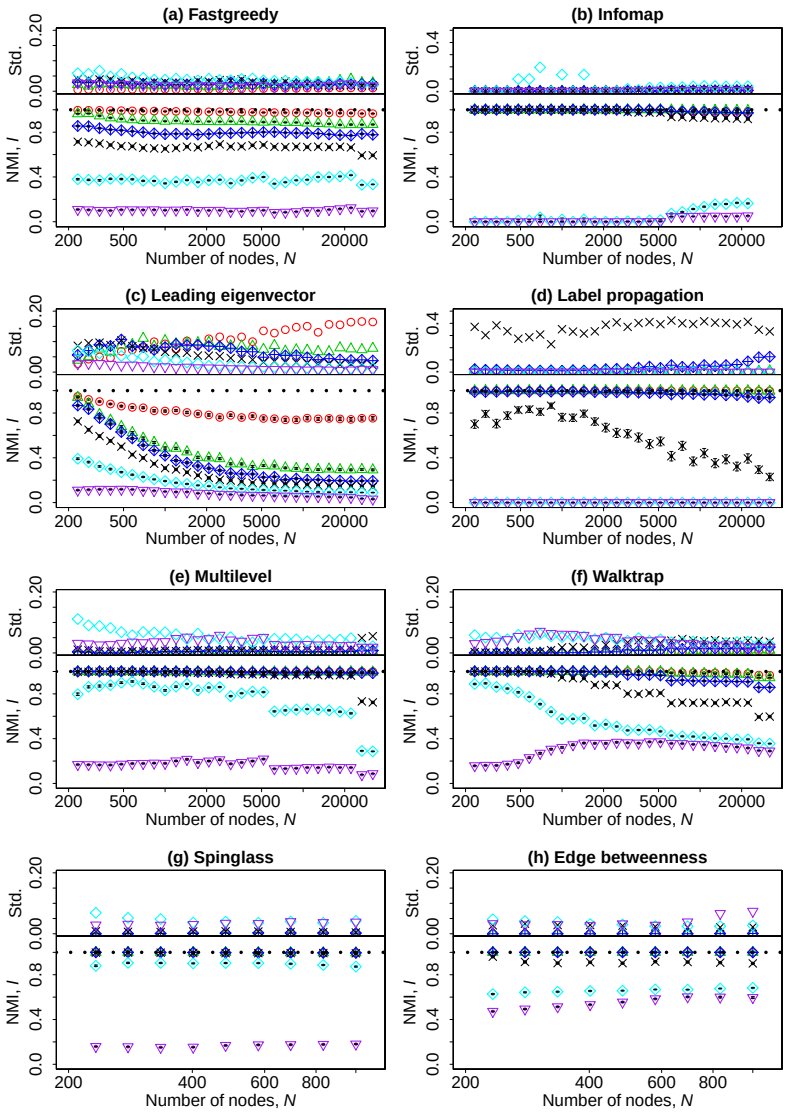
Different colours refer to different values of the mixing parameter: red ($\mu = 0.03$), green ($\mu = 0.18$), blue ($\mu = 0.33$), black ($\mu = 0.48$), cyan ($\mu = 0.63$), and purple ($\mu = 0.75$). Please notice that the vertical axis on the subfigures might have different scale ranges. The horizontal black dotted line corresponds to $I = 1$. Due to the computing speed, Spinglass and Edge betweenness algorithms have been tested only on networks with $N \leq 1000$, and Infomap algorithm has been tested on networks with $N \leq 22186$. The other parameters are described in the main text.

Figure 34: (lower row) The mean value of I_{sum} dependent on the number of nodes N in the benchmark graphs on a *linear-log* scale. (upper row) The standard deviation of I_{sum} dependent on N on a *linear-log* scale.



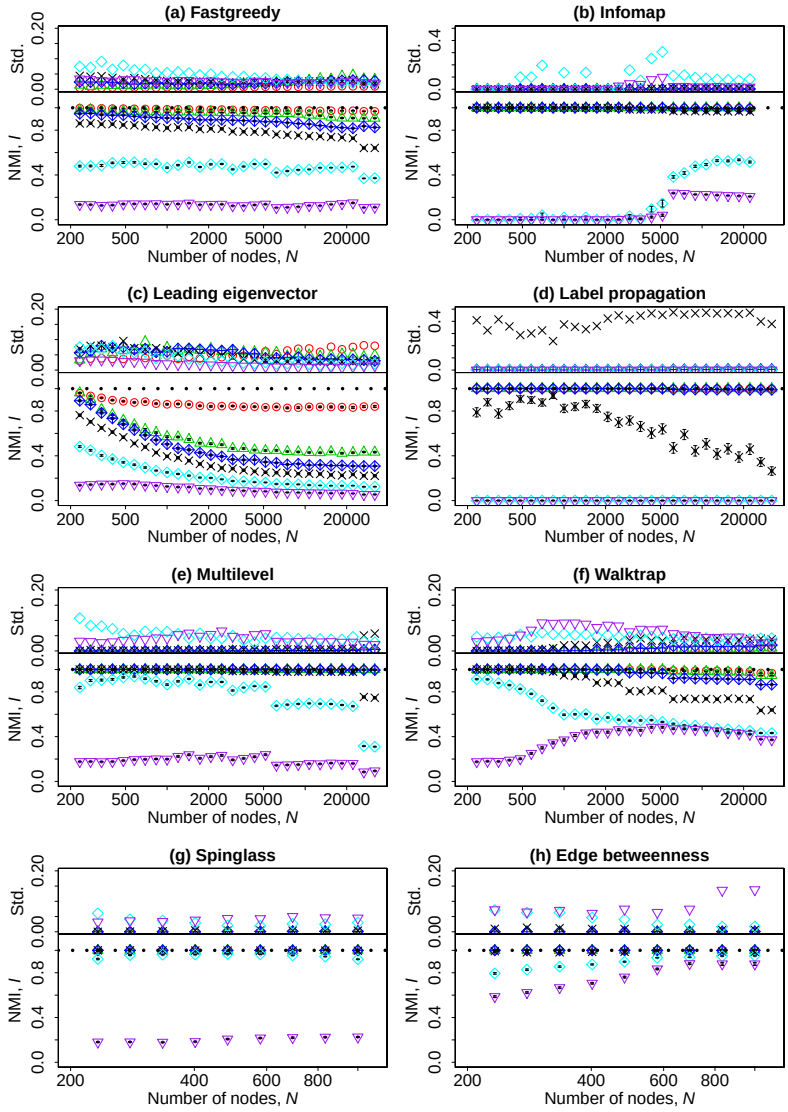
Different colours refer to different values of the mixing parameter: red ($\mu = 0.03$), green ($\mu = 0.18$), blue ($\mu = 0.33$), black ($\mu = 0.48$), cyan ($\mu = 0.63$), and purple ($\mu = 0.75$). Please notice that the vertical axis on the subfigures might have different scale ranges. The horizontal black dotted line corresponds to $I = 1$. Due to the computing speed, Spinglass and Edge betweenness algorithms have been tested only on networks with $N \leq 1000$, and Infomap algorithm has been tested on networks with $N \leq 22186$. The other parameters are described in the main text.

Figure 35: (lower row) The mean value of I_{sqrt} dependent on the number of nodes N in the benchmark graphs on a *linear-log* scale. (upper row) The standard deviation of I_{sqrt} dependent on N on a *linear-log* scale.



Different colours refer to different values of the mixing parameter: red ($\mu = 0.03$), green ($\mu = 0.18$), blue ($\mu = 0.33$), black ($\mu = 0.48$), cyan ($\mu = 0.63$), and purple ($\mu = 0.75$). Please notice that the vertical axis on the subfigures might have different scale ranges. The horizontal black dotted line corresponds to $I = 1$. Due to the computing speed, Spinglass and Edge betweenness algorithms have been tested only on networks with $N \leq 1000$, and Infomap algorithm has been tested on networks with $N \leq 22186$. The other parameters are described in the main text.

Figure 36: (lower row) The mean value of I_{min} dependent on the number of nodes N in the benchmark graphs on a *linear-log* scale. (upper row) The standard deviation of I_{min} dependent on N on a *linear-log* scale.



Different colours refer to different values of the mixing parameter: red ($\mu = 0.03$), green ($\mu = 0.18$), blue ($\mu = 0.33$), black ($\mu = 0.48$), cyan ($\mu = 0.63$), and purple ($\mu = 0.75$). Please notice that the vertical axis on the subfigures might have different scale ranges. The horizontal black dotted line corresponds to $I = 1$. Due to the computing speed, Spinglass and Edge betweenness algorithms have been tested only on networks with $N \leq 1000$, and Infomap algorithm has been tested on networks with $N \leq 22186$. The other parameters are described in the main text.

References

[1] Danon, L., Diaz-Guilera, A., Duch, J. & Arenas, A. Comparing community structure identification. *Journal of Statistical Mechanics: Theory and Experiment* **2005**, P09008 (2005).

[2] Vinh, N. X., Epps, J. & Bailey, J. Information theoretic measures for clusterings comparison: Variants, properties, normalization and correction for chance. *The Journal of Machine Learning Research* **11**, 2837–2854 (2010).

Curriculum Vitae

Personal details

Zhao Yang

Date of birth: 05 May 1988

Education

01/13 – 02/17	Doctoral program at the University of Zurich, Department of Business Administration, Chair of Marketing and Market Research
01/13 – 02/17	Member of the University Research Priority Program on "Social Networks", University of Zurich
07/13 – 08/13	ICPSR summer program, University of Michigan
09/10 – 08/12	Master of Science in Physics, École polytechnique fédérale de Lausanne
09/06 – 08/10	Bachelor of Science in Physics, Joint Bachelor's Degree Program between Université Claude Bernard Lyon 1 (09/08 – 08/10), and Wuhan University (09/06 – 08/08)

Professional experience

01/13 – 02/17	Teaching assistant at the Chair of Marketing and Market Research, University of Zurich